PROTEORHODOPSIN MUTANTS WITH IMPROVED OPTICAL CHARACTERISTICS

This application claims the benefit of Provisional Application 60/429,518, filed November 26, 2002.

FIELD OF INVENTION

This present invention relates to the composition and use of proteorhodopsin mutants with improved optical characteristics. Particularly, the proteorhodopsin mutants have a mutation at a conserved amino acid residue of a naturally occurring proteorhodopsin variant or its equivalent sharing high identity.

BACKGROUND OF THE INVENTION

Proteorhodopsins are integral membrane proteins; they are isolated from uncultivated marine eubacteria and function as light-driven proton pumps. Upon absorption of light by the all-trans-retinal co-factor, proteorhodopsin goes through a photocycle with a number of intermediates. It is believed that upon excitation of the proteorhodopsin molecule by light stimulation, a proteorhodopsin/retinal complex is excited to an unstable intermediate energy state. Proteorhodopsin progresses through a series of unstable energy states that can vary in terms of energy plateaus or intermediates, e.g., an "M-like state" or "M-state", a "K-like state" or "K-state", an "N-like state" or "N-state", or an "O-like state" or "O-state". Subsequently, the complex reverts to a more stable basal state concomitant with transportation of a proton.

Proteorhodopsins are distantly related to bacteriorhodopsin from *Halobacterium salinarium* (22-24% sequence similarity). Hampp (*Appl. Microbiol. Biotechnol.* 53:633-9, 2000a) reviews the structure and function of bacteriorhodopsin, and its technical applications. Hampp (*Chem. Rev.* 100:1755-76, 2000b) reviews the technical application of bacteriorhodopsin.

Proteorhodopsin and bacteriorhodopsin have some shared characteristics, but also have clearly different properties. Proteorhodopsins are more advantageous to use in some technical applications than bacteriorhodopsins because of the ease of expressing and producing proteorhodopsins. However, the conditions where the proteorhodopsins can be used in different applications are limited because wild-type proteorhodopsins exist in two distinct spectral forms depending on the extra-cellular pH. A basic form, which is a spectral form at a higher pH, is able to achieve an M-state of excitation and transport a proton upon

exposure to an optical stimulation. An acidic form, which is a spectral form at a lower pH, is unable to exhibit the M-state of excitation and does not transport a proton upon exposure to an optical stimulation.

The properties of the two distinct pH-dependent spectral forms of the Bac31A8 proteorhodopsin have been characterized to some extent (Dioumaev, *et al.*, *Biochem.* 41:5348-58, 2002; Krebs, *et al.*, *BMC Physiol.* 2:1-8, 2002; Fredrich, et al., J. Mol. Biol. 321, 821-838, 2002). The D97 residue in the Bac31A8 proteorhodopsin was previously identified (Dioumaev, et al., 2002) as being part of the titratable group(s) involved in the pH dependent change in spectral and photochemical properties. However, the Bac31A8 D97N mutant protein appears only to exist as a single spectral form, the acidic form. An analysis of the photocycle intermediates of the Bac31A8 proteorhodopsin at different pH values showed that only the high pH ("basic") form exhibits the photocycle wherein protons are pumped across the membrane. Hence, this Bac31A8 D97N mutant is not very useful for most applications because the protein is unable to pump protons and form an M-state.

Béjà, *et al.* (*Science* 289:1902-6, 2000) disclose the cloning of a proteorhodopsin gene from an uncultivated member of the marine γ-proteobacteria (i.e., the "SAR86" group). The proteorhodopsin was functionally expressed in *E.coli* and bound all-trans-retinal to form an active light-driven proton pump.

Béjà, *et al.* (*Nature* 411:786-9, 2001) disclose the cloning of over twenty variant proteorhodopsin genes from various sources. The proteorhodopsin variants appear to belong to an extensive family of globally distributed proteorhodopsin variants that maximally absorb light at different wavelengths.

WO 01/83701 discloses specific proteorhodopsin gene and protein sequences retrieved from naturally occurring bacteria; the reference also discloses the use of these proteorhodopsin variants in a light-driven energy generation system.

Dioumaev, *et al.* (*Biochem.* 41:5348-58); Krebs, et al. (*BMC Physiol.* 2:1-8, 2002); and Friedrich, *et al* (*J. Mol. Biol.* 321, 821-38, 2002) disclose the properties of two distinct pH-dependent spectral forms of the Bac31A8 proteorhodopsin. Dioumaev, *et al.* also disclose that essentially only the acidic form is present in the Bac31A8 D97N mutant since the N residue is non-protonatable. Further, Dioumaev, *et al.* disclose that the D97E mutant causes minor changes in the absorbance maximum of the acidic and basic forms. An E108Q mutant causes the decay of the M-like state intermediate to be a hundred fold slower. Both the D97E and E108Q mutants have pH titration similar to that of the wild-type protein.

Varo, *et al.* (*Bioophysical J.*, 84:1202-1207 (2003)) describe the results of a thorough analysis of the photocyle of the wildtype Bac31A8 proteorhodopsin; the spectral properties and lifetimes of different intermediates in the photocycles are characterized.

WO 02/10207 discloses proton-translocating retinal protein, such as a *Halobacterium salinarim* bacteriorhodopsin, in which one or more positions of the amino acids that participate in proton-translocation, from the group of amino acid residues D38, R82, D85, D96, D102, D104, E194 and E204 are modified; such proton-translocating retinal proteins have a slower photocycle in comparison to with the wild-type proteins.

SUMMARY OF THE INVENTION

The present invention is directed to a proteorhodopsin mutant having improved optical characteristics; the mutant comprises a mutation in a conserved residue such as a histidine residue and/or an arginine residue of a proteorhodopsin variant. One improved optical characteristic of the mutant is having a lower pH (pK_{rh}), at which equal concentrations of the acidic and basic spectral form of the proteorhodopsin molecules are present. Another improved optical characteristic of the mutant is having a smaller difference in maximum absorption wavelength between the basic and the acidic form. The present invention also provides proteorhodopsin mutants comprising specific amino acid sequences.

A conserved histidine residue is at, for example, amino acid position 75 of Bac31A8, or position 77 of Hot75m1, or its equivalent position of a proteorhodopsin variant. A conserved arginine residue is at, for example, amino acid position 94 of Bac31A8, or position 96 of Hot75m1, or its equivalent position of a proteorhodopsin variant. Preferred mutations are substituting histidine with asparagine, glutamine, lysine, tryptophan, aspartic acid, or glutamic acid; and substituting arginine to alanine, glutamic acid or glutamine.

The present invention also provides an isolated nucleic acid sequence encoding the proteorhodopsin mutant. The present invention further provides a method for preparing a proteorhodopsin mutant having improved optical characteristics.

DESCRIPTION OF THE FIGURES

FIGURE 1-1 depicts the amino acid and nucleotide sequences of Hot75m1 (SEQ ID NOs: 1 and 2).

FIGURE 1-2 depicts the amino acid and nucleotide sequences of Bac31A8 (SEQ ID NOs: 3 and 4).

FIGURE 1-3 depicts the amino acid and nucleotide sequences of Bac40E8 (SEQ ID NOs: 5 and 6).

FIGURE 1-4 depicts the amino acid and nucleotide sequences of Bac64A5 (SEQ ID NOs: 7 and 8).

FIGURE 1-5 depicts the amino acid and nucleotide sequences of Hot0m1 (SEQ ID NOs: 9 and 10).

FIGURE 1-6 depicts the amino acid and nucleotide sequences of Hot75m3 (SEQ ID NOs: 11 and 12).

FIGURE 1-7 depicts the amino acid and nucleotide sequences of Hot75m4 (SEQ ID NOs: 13 and 14).

FIGURE 1-8 depicts the amino acid and nucleotide sequences of Hot75m8 (SEQ ID NOs: 15 and 16).

FIGURE 1-9 depicts the amino acid and nucleotide sequences of MB0m1 (SEQ ID NOs: 17 and 18).

FIGURE 1-10 depicts the amino acid and nucleotide sequences of MB0m2 (SEQ ID NOs: 19 and 20).

FIGURE 1-11 depicts the amino acid and nucleotide sequences of MB20m2 (SEQ ID NOs: 21 and 22).

FIGURE 1-12 depicts the amino acid and nucleotide sequences of MB20m5 (SEQ ID NOs: 23 and 24).

FIGURE 1-13 depicts the amino acid and nucleotide sequences of MB20m12 (SEQ ID NOs: 25 and 26).

FIGURE 1-14 depicts the amino acid and nucleotide sequences of MB40m1 (SEQ ID NOs: 27 and 28).

FIGURE 1-15 depicts the amino acid and nucleotide sequences of MB40m5 (SEQ ID NOs: 29 and 30).

FIGURE 1-16 depicts the amino acid and nucleotide sequences of MB40m12 (SEQ ID NOs: 31 and 32).

FIGURE 1-17 depicts the amino acid and nucleotide sequences of MB100m5 (SEQ ID NOs: 33 and 34).

FIGURE 1-18 depicts the amino acid and nucleotide sequences of MB100m7 (SEQ ID NOs: 35 and 36).

FIGURE 1-19 depicts the amino acid and nucleotide sequences of MB100m9 (SEQ ID NOs: 37 and 38).

FIGURE 1-20 depicts the amino acid and nucleotide sequences of MB100m10 (SEQ ID NOs: 39 and 40).

FIGURE 1-21 depicts the amino acid and nucleotide sequences of PalB1 (SEQ ID NOs: 41 and 42).

FIGURE 1-22 depicts the amino acid and nucleotide sequences of PalB2 (SEQ ID NOs: 43 and 44).

FIGURE 1-23 depicts the amino acid and nucleotide sequences of PalB5 (SEQ ID NOs: 45 and 46).

FIGURE 1-24 depicts the amino acid and nucleotide sequences of PalB7 (SEQ ID NOs: 47 and 48).

FIGURE 1-25 depicts the amino acid and nucleotide sequences of PalB6 (SEQ ID NOs: 49 and 50).

FIGURE 1-26 depicts the amino acid and nucleotide sequences of PalB8 (SEQ ID NOs: 51 and 52).

FIGURE 1-27 depicts the amino acid and nucleotide sequences of PalE1 (SEQ ID NOs: 53 and 54).

FIGURE 1-28 depicts the amino acid and nucleotide sequences of PalE6 (SEQ ID NOs: 55 and 56).

FIGURE 1-29 depicts the amino acid and nucleotide sequences of PalE7 (SEQ ID NOs: 57 and 58).

FIGURE 1-30 depicts the amino acid and nucleotide sequences of MED26 (SEQ ID NOs: 59 and 60).

FIGURE 1-31 depicts the amino acid and nucleotide sequences of MED27 (SEQ ID NOs: 61 and 62).

FIGURE 1-32 depicts the amino acid and nucleotide sequences of MED36 (SEQ ID NOs: 63 and 64).

FIGURE 1-33 depicts the amino acid and nucleotide sequences of MED101 (SEQ ID NOs: 65 and 66).

FIGURE 1-34 depicts the amino acid and nucleotide sequences of MED102 (SEQ ID NOs: 67 and 68).

FIGURE 1-35 depicts the amino acid and nucleotide sequences of MED106 (SEQ ID NOs: 69 and 70).

FIGURE 1-36 depicts the amino acid and nucleotide sequences of MED25 (SEQ ID NOs: 71 and 72).

FIGURE 1-37 depicts the amino acid and nucleotide sequences of MED202 (SEQ ID NOs: 73 and 74).

FIGURE 1-38 depicts the amino acid and nucleotide sequences of MED204 (SEQ ID NOs: 75 and 76).

FIGURE 1-39 depicts the amino acid and nucleotide sequences of MED208 (SEQ ID NOs: 77 and 78).

FIGURE 1-40 depicts the amino acid and nucleotide sequences of REDA9 (SEQ ID NOs: 79 and 80).

FIGURE 1-41 depicts the amino acid and nucleotide sequences of REDB9 (SEQ ID NOs: 81 and 82).

FIGURE 1-42 depicts the amino acid and nucleotide sequences of REDF9 (SEQ ID NOs: 83 and 84).

FIGURE 1-43 depicts the amino acid and nucleotide sequences of RED19 (SEQ ID NOs: 85 and 86).

FIGURE 1-44 depicts the amino acid and nucleotide sequences of RED2 (SEQ ID NOs: 87 and 88).

FIGURE 1-45 depicts the amino acid and nucleotide sequences of RED23 (SEQ ID NOs: 89 and 90).

FIGURE 1-46 depicts the amino acid and nucleotide sequences of RED27 (SEQ ID NOs: 91 and 92).

FIGURE 1-47 depicts the amino acid and nucleotide sequences of RED30 (SEQ ID NOs: 93 and 94).

FIGURE 1-48 depicts the amino acid and nucleotide sequences of RED4 (SEQ ID NOs: 95 and 96).

FIGURE 1-49 depicts the amino acid and nucleotide sequences of RED5 (SEQ ID NOs: 97 and 98).

FIGURE 1-50 depicts the amino acid and nucleotide sequences of REDr6a5a14 (SEQ ID NOs: 99 and 100).

FIGURE 1-51 depicts the amino acid and nucleotide sequences of REDr6a5a6 (SEQ ID NOs: 101 and 102).

FIGURE 1-52 depicts the amino acid and nucleotide sequences of REDr7_1_4 (SEQ ID NOs: 103 and 104).

FIGURE 1-53 depicts the amino acid and nucleotide sequences of REDs3_7 (SEQ ID NOs: 105 and 106).

FIGURE 1-54 depicts the amino acid and nucleotide sequences of REDr7_1_15 (SEQ ID NOs: 107 and 108).

FIGURE 1-55 depicts the amino acid and nucleotide sequences of REDs3_15 (SEQ ID NOs: 109 and 110).

FIGURE 1-56 depicts the amino acid and nucleotide sequences of medA15r8ex6 (SEQ ID NOs: 111 and 112).

FIGURE 1-57 depicts the amino acid and nucleotide sequences of REDr7_1_16 (SEQ ID NOs: 113 and 114).

FIGURE 1-58 depicts the amino acid and nucleotide sequences of medA15r11b9 (SEQ ID NOs: 115 and 116).

FIGURE 1-59 depicts the amino acid and nucleotide sequences of medA15r9b5 (SEQ ID NOs: 117 and 118).

FIGURE 1-60 depicts the amino acid and nucleotide sequences of medA15r8b3 (SEQ ID NOs: 119 and 120).

FIGURE 1-61 depicts the amino acid and nucleotide sequences of medA15r11b3 (SEQ ID NOs: 121 and 122).

FIGURE 1-62 depicts the amino acid and nucleotide sequences of medA15_r8_1 (SEQ ID NOs: 123 and 124).

FIGURE 1-63 depicts the amino acid and nucleotide sequences of medA17R9_1 (SEQ ID NOs: 125 and 126).

FIGURE 1-64 depicts the amino acid and nucleotide sequences of medA15r8b9 (SEQ ID NOs: 127 and 128).

FIGURE 1-65 depicts the amino acid and nucleotide sequences of medA19_R8_16 (SEQ ID NOs: 129 and 130).

FIGURE 1-66 depicts the amino acid and nucleotide sequences of medA19_R8_19 (SEQ ID NOs: 131 and 132).

FIGURE 1-67 depicts the amino acid and nucleotide sequences of medA17_R8_6 (SEQ ID NOs: 133 and 134).

FIGURE 1-68 depicts the amino acid and nucleotide sequences of medA15r9b7 (SEQ ID NOs: 135 and 136).

FIGURE 1-69 depicts the amino acid and nucleotide sequences of medA15_R8_3 (SEQ ID NOs: 137 and 138).

FIGURE 1-70 depicts the amino acid and nucleotide sequences of medA15r10b5 (SEQ ID NOs: 139 and 140).

FIGURE 1-71 depicts the amino acid and nucleotide sequences of medA19_r9_9 (SEQ ID NOs: 141 and 142).

FIGURE 1-72 depicts the amino acid and nucleotide sequences of medA15_r8ex7 (SEQ ID NOs: 143 and 144).

FIGURE 1-73 depicts the amino acid and nucleotide sequences of medA19_R8_20 (SEQ ID NOs: 145 and 146).

FIGURE 1-74 depicts the amino acid and nucleotide sequences of medA15_R8ex9 (SEQ ID NOs: 147 and 148).

FIGURE 1-75 depicts the amino acid and nucleotide sequences of medA15_r9_3 (SEQ ID NOs: 149 and 150).

FIGURE 1-76 depicts the amino acid and nucleotide sequences of medA17_r8_15 (SEQ ID NOs: 151 and 152).

FIGURE 1-77 depicts the amino acid and nucleotide sequences of medA17_r8_11 (SEQ ID NOs: 153 and 154).

FIGURE 1-78 depicts the amino acid and nucleotide sequences of medA15r8b8 (SEQ ID NOs: 155 and 156).

FIGURE 1-79 depicts the amino acid and nucleotide sequences of medA15r8ex4 (SEQ ID NOs: 157 and 158).

FIGURE 1-80 depicts the amino acid and nucleotide sequences of ANT32C12 PR (SEQ ID NOs: 159 and 160).

FIGURE 1-81 depicts the amino acid and nucleotide sequences of HOT2C01 PR (SEQ ID NOs: 161 and 162).

Figures 2-1 to 2-11 show the amino acid and nucleotide sequences of proteorhodopsin mutants Bac31A8 H75K (SEQ ID NOs: 163 and 164), Bac31A8 H75N (SEQ ID NOs: 165 and 166), Bac31A8 H75Q (SEQID NOs: 167 and 168), Hot75ml H77K (SEQ ID NOs: 169

and 170), Hot75ml H77N (SEQ ID NOs: 171 and 172), Hot75ml H77Q (SEQ ID NOs: 173 and 174), Hot75ml H77E (SEQ ID NOs: 175 and 176), Hot75ml H77W (SEQ ID NOs: 177 and 178), Hot75ml R96A (SEQ ID NOs: 179 and 180), Hot75ml R96E (SEQ ID NOs: 181 and 182), and Hot75ml R96Q (SEQ ID NOs: 183 and 184).

FIGURES 3-1 to 3-8 depict an alignment of the amino acid sequences of 81 natural proteorhodopsin variants. The bold "**H**" indicates the position of a conserved histidine, which corresponds to H75 of Bac31A8. The bold "**R**" indicates the position of a conserved arginine, which corresponds to R94 of Bac31A8.

FIGURE 4 depicts the map of plasmid pTrcHis2-Hot75m1.

FIGURE 5 shows the titration curves of absorption spectra of mutant Hot75m1 H77Q and wildtype (WT) Hot75m1 at pH 4.9, 6.0, 6.9, 8.1 and 9.1.

FIGURE 6 shows the fraction of acidic form of Hot75ml and Hot75ml H77Q at various pH's.

FIGURE 7-1 shows the extracellular pH changes with cells expressing Bac31A8 wild-type, with and without illumination.

FIGURE 7-2 shows the extracellular pH changes with cells expressing a LacZ control protein, with and without illumination.

FIGURE 7-3 shows the effect of extracellular pH on the proton-pumping rate for Bac31A8 wild-type and the H75N mutant.

DETAILED DESCRIPTION OF THE INVENTION

Definitions

As used in the specification and claims, the singular "a", "an" and "the" include the plural references unless the context clearly dictates otherwise. For example, the term polypeptide may include a plurality of polypeptides.

The term "derived" shall encompass derivation by information alone. For example, an amino acid sequence can be derived from a wild-type protein by using the information of the known amino acid sequence of the wild-type protein to chemically synthesize the amino acid sequence.

The term "proteorhodopsin variants" encompasses various naturally occurring proteorhodopsins and their homologues, either known or unknown, which are able to undergo a photocycle containing an "M-state" or "M- like state." FIGURES 1-1 to 1-81 provide the nucleotide and amino acid sequences of 81 proteorhodopsin variants; these proteorhodopsin

variants show between 53 to 99.6% amino acid sequence identity to each other. The term "proteorhodopsin variant" includes proteins having an amino acid sequence that is at least 53%, preferably 90%, more preferably 95%, most preferably 97 or 98%, identical to that of any proteorhodopsin variant described in Figures 1-1 to 1-81. The term "proteorhodopsin variant" also encompasses fused proteorhodopsin constructed by fusing the amino acid sequences from two or more different naturally found proteorhodopsins, so that each amino acid sequence occupies an equivalent position in the resulting proteorhodopsin (even if these fused proteorhodopsin are not naturally occurring).

The term "proteorhodopsin mutant" for the purpose of this application refers to a proteorhodopsin variant comprising one or more mutations that substitutes one or more amino acid residues and/or nucleotides by different amino acids and/or nucleotide sequences.

The term "basal state" or "B-state" or B-like state" refers to the basal state of the photocycle of a proteorhodopsin molecule without light excitation; the basal absorption maxima of proteorhodopsin variants are in general between 480 nm and 530 nm, often between 488 and 526 nm.

The term "M-state" or "M-like state" refers to an excited spectral state in a photocycle as compared with the basal state; the absorption maxima of the M-state of proteorhodopsin variants in general are between 350 nm and 450 nm, often about 410 nm. The M-state is distinguished from other identified spectral states, the K-, N- and O-like states, which all have red-shifted absorbtion spectra (e.g. >530 nm) compared with the basal state.

The term "pK_{rh}" for the purpose of this application, refers to the pH at which equal concentrations of the acidic and basic spectral forms of the proteorhodopsin molecule are present. Applicants derive this term from pK_a, which is a term that identifies the pH at which equal concentrations of the acidic and basic forms of the molecule are present. If the pK_{rh} of that particular proteorhodopsin is determined to be a specific value, that specific value is the pH where the basic and acidic form are present in equal concentrations. Shifting the pH of the environment relative to the pK_{rh} will increase the concentration of the basic or acidic form present, depending upon the direction (more acidic or more basic) and the pH units shifted.

The term "wavelength maximum" for the purpose of this application, is the wavelength of maximum absorbance for proteorhodopsin at a specific pH.

Abbreviations

IPTG, Isopropyl β-D-thiogalactopyranoside.

MES, 2-(N-morpholino)ethanesulfonic acid.

MOPS, (3-N-morpholino)propanesulfonic acid.

TAPS, N-([*tris*(hydroxymethyl)methyl]amino)propanesulfonic acid.

CHES, 2-(N-cyclohexlamino)ethanesulfonic acid.

EDTA, Ethylenediaminetetraacetate.

Deposits

The *E. coli* containing the Bac31A8 clone (assigned ATCC No. PTA-3083), the *E. coli* containing the Bac40E8 clone (assigned ATCC No. PTA-3082), the *E. coli* containing the Bac41B4 clone (assigned ATCC No. PTA-3080), and the *E. coli* containing the Bac64A5 clone (assigned ATCC No. PTA-3081) were deposited with the ATCC Patent Depository (10801 University Blvd., Manassas, VA 20110, U.S.A.) on February 21, 2001.

The *E. coli* containing plasmid PalE6 (assigned ATCC No. PTA-3250), the *E. coli* containing the plasmid Hot0m1 (assigned ATCC No. PTA-3251), and the *E. coli* containing the plasmid Hot75m4 (assigned ATCC No. PTA-3252) were deposited with the ATCC Patent Depository (10801 University Blvd., Manassas, VA 20110, U.S.A.) on March 30, 2001.

Present Invention

The present invention is directed to a proteorhodopsin mutant having an improved optical characteristic. The mutant comprises a mutation in a conserved amino acid residue of a proteorhodopsin variant, which causes the spectral shifts. The improved optical characteristics include having a lower pK_{rh} or a smaller difference in maximum absorption wavelength between the basic and the acidic form, in comparison with the proteorhodopsin variant from which the mutant is derived.

Proteorhodopsin and its Photochemical Properties

Proteorhodopsin is a trans-membrane protein with a structure of seven lipid membrane-spanning α-helices which form a generally cylinder shaped channel. When folded correctly and supplied with all-trans-retinal, the seven α-helices of proteorhodpsin are arranged as a cage surrounding the all-trans-retinal. A properly folded proteorhodopsin has the property of being able to bind all-trans-retinal and undergo a photocycle wherein a proton is transported. The source of all-trans-retinal includes chromophore retinal and chemical derivatives of all-trans-retinal. Chemical derivatives of retinal include, but are not limited to,

3-methyl-5-(1-pyryl)-2E,4E-pentadienal, 3,7-dimethyl-9-(1-pyryl)-2E,4E,6E,8E-nonatetraenal, all-trans-9-(4-azido-2,3,5,6-tetrafluorophenyl)-3,7-dimethyl-2,4,6,8-nonatetraenal, 2,3-dehydro-4-oxoretinal, and like compounds. Proteorhodopsin is a light-activated proton pump. Proteorhodopsin binds all-trans-retinal to form a pigment that absorbs in the visible wavelength range of light.

An all-trans-retinal covalently attached to a conserved lysine (K231 in Bac31A8 and K234 in Hot75M1) by a Schiff-base linkage contributes to the visible light chromophore of proteorhodopsin. The absorbance of light energy by this chromophore is converted through a photocycle into mechanical energy that pumps a proton from the interior to the exterior of the cellular host through the all-trans-retinal binding pocket. The resulting proton imbalance is then used by the cell as chemical energy.

Proteorhodopsin has two distinct pH-dependent spectral forms: a basic form and an acidic form. The basic form can undergo a photocycle that includes the excited M-state (or the M intermediate) and pump a proton, H^+, across the proteorhodopsin-containing membrane, from inside the cell to outside the cell. The basic form has a rapid photocycle and is able to transport protons out of a cell or cell vesicle. When the proteorhodopsin is present in a basic form and is stimulated by light, it first proceeds into an excited energy state, e.g. an M-state, then it proceeds back into the more basal energy or resting state, where upon it transports a proton. The charge or proton is transported thereby, pumping a proton through the membrane and out of the cell during this process. (See Dioumaev, *et al.* and Krebs, *et al.*)

The acidic form of proteorhodopsin is unable to undergo a photocycle that includes the M-state, and is unable to pump protons. Spectroscopic titration curves, plotting wavelength against absorbance in buffers of different pH, indicate that the concentration of the acidic form and the basic form of some wild-type proteorhodopsins are about equal at a pH range about 7-8 (i.e. pK_{rh} is about 7- 8). For example, Hot75m1 in intact cells has a pK_{rh} value of about 8.2. Hot75m1 exhibits a shift in the absorption maxima of about 45 nm when the pH is changed from 4.9 to 9.1. Because only one form, the basic form, gives a productive photocycle, the formation of different spectral forms limits the conditions where the proteorhodopsins can be used for certain applications.

Different naturally occurring proteorhodopsin variants have different absoption maxima wavelengths. The absorption maxima wavelengths when the proteorhodopsin is in an acidic form range from about 534 to about 570 nm. The maximum wavelength when the proteorhodopsin is in a basic form falls within two groups: those that range from about 488 to

about 494 nm, and those that range from about 516 to about 528 nm. Different naturally occurring proteorhodopsin variants also have different pK_{rh} values, ranging from about 7.1 to about 8.6.

Without wishing to be bound by any theory, it is believed that the photochemical property of proteorhodopsin can be altered by, either one or a combination of, the pH, presence of an amphipathic molecule (such as a surfactant detergent), temperature of the environment, the presence of chemical additives in the environment (such as azide or glycerol or the like), and the water content of the medium containing the photeorhodopsin.

Improved Optical Property

As described above, a useful and productive form of proteorhodopsin is the basic form. To achieve a predominantly basic spectral form of a proteorhodopsin population with a pK_{rh} of 8.0, a two or more pH unit shift to pH 10.0 or more is contemplated to exhibit an increased relative amount of the basic form versus the acidic form. However, such a basic environment is not an optimal pH for proteorhodopsins, since proteins in general tend to denature at such a high pH. Therefore, a proteorhodopsin mutant having a low pK_{rh} value is desirable.

The proteorhodopsin mutant of the present invention preferably has an altered pK_{rh} value that is lower (more acidic) than that of the naturally occurring proteorhodopsin. In one embodiment, the proteorhodopsin mutant has a pK_{rh} value that is lower than neutral pH. This is because at near neutral pH, the polypeptide is more stable and has a longer shelf life. To allow the proteorhodopsin to undergo a photocycle wherein a proton is transported at near neutral pH, the pK_{rh} value of the proteorhodopsin mutant should be reduced as much as possible so that the basic form can predominate at near a neutral pH range, for example, about 5.5-9.0, or about 6.0-9.0, or about 7.1 - 8.6.

The pK_{rh} value of the proteorhodopsin mutant is reduced in comparison with that of the proteorhodopsin variant from which the mutant is derived. The pK_{rh} value is reduced by at least 1.0, preferably, 2.0, and more preferably, 3.0 pH units.

In general, the value of the pK_{rh} of proteorhodopsin mutant is reduced to lower than 7.0, preferably 6.0, more preferably 5.0, more preferably 4.0, and even more preferably 3.0.

A proteorhodopsin mutant in general has a broad range of pH in which the basic form predominates. The range of the pH values where the basic form of proteorhodopsin predominates is increased in the mutant by 1.0, preferably by 2.0, and more preferably by 3.0,

when compared with the proteorhodopsin variant that the mutant is derived from. Preferably, the pK_{rh} value of the proteorhodopsin mutant is low enough such that essentially only the basic form is present in the useful pH range.

In the proteorhodopsin mutant, the lower limit of the pH range where the basic form predominates is extended to at least pH 7.0, preferably 6.0, more preferably 5.0, more preferably 4.0, or 3.0. Alternatively, the proteorhodopsin mutant predominates in the basic form under all pH conditions.

In another embodiment of the invention, the proteorhodopsin mutant has a smaller difference in maximum absorption wavelength between the basic and the acidic form, in comparison with the proteorhodopsin variant from which the mutant is derived. Such mutants, of which the basic and acidic absorption maxima are closer to each other, are useful in applications where the environmental pH is close to the pK_{rh} value. In such applications, both spectral forms will be present. A large difference in the absorption maxima of the acidic and basic forms will result in a broad composite absorption spectra (broader spectral width) because the acidic and basic absorption spectra are superimposed. When a mutant proteorhodopsin has a smaller difference in the absorption maxima of the acidic and basic forms, the composite absorption spectra will be narrower (smaller peak width).

The proteorhodopsin mutant has a difference in maximum absorption wavelength between the basic and the acidic form of less than 25 nm, preferably less than 20 nm, more preferably less than 15 nm, and most preferably, less than 10 nm.

Proteorhodopsin variants

A proteorhodopsin variant that the mutant is derived from can be any naturally occurring proteorhodopsin. Proteorhodopsin variants, and nucleic acid sequences encoding thereof, have been obtained from naturally occurring members of the domain bacteria. Such members include marine bacteria, such as bacteria from the SAR86 group. Proteorhodopsin variants useful in the present invention include those derived from marine bacteria, as well as those derived or obtained from non-marine bacteria. There are many variant forms of proteorhodopsin; all of which can be used for the present invention.

The amino acid sequences of 29 proteorhodopsin variants from various sources are shown in Figures 1-1 to 1-29 (Baja, *et al.*, *Nature* 411:786-9 (2001)). These include Hot75m1 (SEQ ID NO: 1), Bac31A8 (SEQ ID NO: 3), Bac40E8 (SEQ ID NO: 5), Bac41B4 (SEQ ID NO: 7), Bac64A5 (SEQ ID NO: 9), Hot0m1 (SEQ ID NO: 11), Hot75m3 (SEQ ID

NO: 13), Hot75m4 (SEQ ID NO: 15), Hot75m8 (SEQ ID NO: 17), MB0m1 (SEQ ID NO: 19), MB0m2 (SEQ ID NO: 21), MB20m2 (SEQ ID NO: 23), MB20m5 (SEQ ID NO: 25), MB20m12 (SEQ ID NO: 27), MB40m1 (SEQ ID NO: 29), MB40m5 (SEQ ID NO: 31), MB100m5 (SEQ ID NO: 33), MB100m7 (SEQ ID NO: 35), MB100m9 (SEQ ID NO: 37), MB100m10 (SEQ ID NO: 39), PalB1 (SEQ ID NO: 41), PalB2 (SEQ ID NO: 43), PalB5 (SEQ ID NO: 45), PalB7 (SEQ ID NO: 47), PalB6 (SEQ ID NO: 49), PalB8 (SEQ ID NO: 51), PalE1 (SEQ ID NO: 53), PalE6 (SEQ ID NO: 55), and PalE7 (SEQ ID NO: 57).

The amino acid sequences of another 22 proteorhodopsin variants from the Mediterranean Sea and Red Sea are shown in Figures 1-30 to 1-51 (see Man, *et al.*, *EMBO J.*, 22:1725-1731 (2003)). These include MED 26 (SEQ ID NO: 59), MED 27 (SEQ ID NO: 61)), MED36 (SEQ ID NO: 63), MED101 (SEQ ID NO: 65), MED102 (SEQ ID NO: 67), MED106 (SEQ ID NOs: 69). MED25 (SEQ ID NO: 71), MED202 (SEQ ID NOs: 73), MED204 (SEQ ID NO: 75). MED208 (SEQ ID NO: 77), REDA9 (SEQ ID NOs: 79), REDB9 (SEQ ID NO: 81), REDF9 (SEQ ID NO: 83), RED19 (SEQ ID NO: 85), RED2 (SEQ ID NO: 87), RED23 (SEQ ID NO: 89), RED27 (SEQ ID NO: 91), RED30 (SEQ ID NO: 93), RED4 (SEQ ID NO: 95), RED5 (SEQ ID NO: 97), REDr6a5a14 (SEQ ID NO: 99), and REDr6a5a6 (SEQ ID NO: 101).

The amino acid sequences of another 28 proteorhodopsin variants from the Mediterranean Sea and Red Sea are shown in Figures 1-52 to 1-79 (see Sabehi, *et al.*, *Environ. Microbiol.*, 5: 842-9 (2003)). These include REDr7_1_4 (SEQ ID NO: 103), REDs3_7 (SEQ ID NO: 105), REDr7_1_15 (SEQ ID NO: 107), REDs3_15 (SEQ ID NO: 109), medA15r8ex6 (SEQ ID NO: 111), REDr7_1_16 (SEQ ID NO: 113), medA15r11b9 (SEQ ID NO: 115), medA15r9b5 (SEQ ID NO: 117), medA15r8b3 (SEQ ID NO: 119), medA15r11b3 (SEQ ID NO: 121), medA15_r8_1 (SEQ ID NO: 123), medA17R9_1 (SEQ ID NO: 125), medA15r8b9 (SEQ ID NO: 127), medA19_R8_16 (SEQ ID NO: 129), medA19_R8_19 (SEQ ID NO: 131), medA17_R8_6 (SEQ ID NO: 133), medA15r9b7 (SEQ ID NO: 135), medA15_R8_3 (SEQ ID NO: 137), medA15r10b5 (SEQ ID NO: 139), medA19_r9_9 (SEQ ID NO: 141), medA15_r8ex7 (SEQ ID NO: 143), medA19_R8_20 (SEQ ID NO: 145), medA15_R8ex9 (SEQ ID NO: 147), medA15_r9_3 (SEQ ID NO: 149), medA17_r8_15 (SEQ ID NO: 151), medA17_r8_11 (SEQ ID NO: 153), medA15r8b8 (SEQ ID NO: 155), and medA15r8ex4 (SEQ ID NO: 157).

The amino acid sequences of another 2 newly isolated proteorhodopsin variants from are shown in Figures 1-80 to 1-81 (see De La Torre, *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 100:

12830-5 (2003)). These include ANT32C12 PR (SEQ ID NO: 159) and HOT2C01 PR (SEQ ID NO: 161).

The proteorhodopsin variant, from which the mutant is derived, can be a naturally occurring proteorhodopsin variant, including but not limited to those of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, and 161, or other proteorhodopsin variants sharing at least 70%, or 80%, or 90%, or 95% amino acid identity with those listed sequences.

The nucleotide and amino acid sequences of the 81 proteorhodopsin variants (Figure 1-1 to 1-81) can be used to prepare mutants for this invention. In addition, the nucleotide and amino acid sequences of the 81 proteorhodopsin variants can be altered by substitutions, additions or deletions to provide functionally equivalent molecules, which are suitable for preparing mutants for this invention. For example, due to the degeneracy of nucleotide coding sequences, other DNA sequences that encode substantially the same amino acid sequence as depicted in Figures 1-1 through 1-81 can be used in the practice of the present invention. The DNA sequence can be altered by a substitution of a different codon that encodes the same or a functionally equivalent amino acid residue within the sequence, thus producing a silent change. For example, an amino acid residue within the sequence can be substituted by another amino acid of a similar polarity, or a similar class. Non-polar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, glycine and methionine. Polar neutral amino acids include serine, threonine, cysteine, tyrosine, asparagine, and glutamine. Positively charged (basic) amino acids include arginine, lysine, and histidine. Negatively charged (acidic) amino acids include aspartic and glutamic acid.

In one embodiment of the invention, a proteorhodopsin mutant comprises the amino acid sequence of Bac31A8 H75K (SEQ ID NO: 163), Bac31A8 H75N (SEQ ID NO: 165), Bac31A8 H75Q (SEQ ID NO: 167), Hot75ml H77K (SEQ ID NO: 169), Hot75ml H77N (SEQ ID NO: 171), Hot75ml H77Q (SEQ ID NO: 173), Hot 75ml H77E (SEQ ID NO: 175), Hot75ml H77W (SEQ ID NO: 177), Hot75ml R96A (SEQ ID NO: 179), Hot75ml R96E (SEQ ID NO: 181) and Hot75ml R96Q (SEQ ID NO: 183).

Proteorhodopsin mutation sites

Naturally occurring proteorhodopsin variants have about 234 to 249 amino acids. Comparing the amino acid sequence of different forms of proteorhodopsins, while contrasting their physical, or chemical properties, can reveal specific target regions that are likely to produce useful mutant proteins, and direct the creation of new mutants with deliberately modified functions.

According to the invention, the sequences determined for proteorhodopsins Bac31A8 and/or Hot75m1 are compared with sequences of known proteorhodopsins (see Figs. 1-1 through 1-81) or newly discovered proteorhodpsins in order to deduce sites for desirable mutations. To do this, the closeness of relation of the proteorhodopsins being compared is first determined.

Closeness of relation can be measured by comparing of amino-acid sequences. There are many methods of aligning protein sequences. Methods defining relatedness are described in Atlas of Protein Sequence and Structure, Margaret O. Dayhoff editor, vol. 5, supplement 2, 1976, National Biomedical Research Foundation, Georgetown University Medical Center, Washington, D.C., p. 3 ff., entitled SEARCH and ALIGN. As known in the art, related proteins can differ in the number of amino acids as well as the identity of each amino acid along the chain. That is, there can be deletions or insertions when two structures are aligned for maximum identity. For example, proteorhodopsin Bac31A8 has only 249 amino acids while proteorhodopsin Hot75m1 has 252 amino acids. Aligning the two sequences shows that Bac31A8 has no residue corresponding to 214 of Hot75M1. Thus, the amino acid sequence of Bac31A8 would appear very different from Hot75m1 unless a gap is recorded between locations 211 and 212 of Bac31A8 (see Figure 2 for alignment). Based on the proper amino acid sequence alignment of Bac31A8 and Hot75m1, one can predict with a high degree of confidence that substituting Q, N or K for H at location 75 of proteorhodopsin Bac31A8 will incur the same altered photochemical and spectrophotometric characteristics as substituting Q, N or K for H at location 77 of proteorhodopsin Hot75M1.

The conserved amino acid D97 (Asp97) in the Bac31A8 proteorhodopsin is the amino acid residue that donates a proton to the retinal Schiff-base during the photocycle. The Bac31A8 D97N proteorhodopsin mutant is locked in the acidic form and is therefore incapable of pumping protons or undergoing a complete photocycle (Dioumaev, *et al.*). The Bac31A8 D97E proteorhodopsin mutant, which has a conserved amino acid replacement, has different wavelength spectra at acidic and basic pH compared with those of the wild-type Bac31A8 proteorhodopsin. However, the Bac31A8 D97E mutant has a similar pH dependent

spectral change and a similar pK_{rh} value compared with that of wild-type Bac31A8 proteorhodopsin (Dioumaev, *et al.*). Thus, a conservative replacement in the Asp97 position causes spectral changes, but not a change in the pK_{rh} value. Since the nature and position of the proton-donating group is important for the proper functioning of the proteorhodopsin photocycle, it is unlikely that changes in D97 of Bac31A8 proteorhodopsin will allow for alteration of the pK_{rh} value without disrupting the pumping of protons or the photocycle.

Conserved amino acid residues that are involved in the relay of protons through the all-trans-retinal binding site, but are not in direct contact with the all-trans-retinal cofactor, are likely to affect the pH dependent spectral shift and the pK_{rh} value of proteorhodopsin, and allow for a continued pumping of protons and therefore a productive photocycle. Applicants have identified such conserved amino acid residues in proteorhodopsins.

The present invention provides a mutant proteorhodopsin wherein said proteorhodopsin has one or more mutations in conserved amino acid residues, wherein said one or more mutations cause said proteorhodopsin to have an altered photochemical property, wherein said mutant proteorhodopsin, when optically stimulated, undergoes a photocycle in which a proton is transported. The proteorhodopsin mutant of this invention has an improved optical characteristic in comparison with the proteorhodopsin variant. The improved optical characteristic is a lower pK_{rh} value or a smaller difference in maximum absorption wavelength between a basic and an acidic form. The present invention has identified a number of conserved amino acid residues that may interact with the proton transported during the photocycle, thus altering the pH at which the basic form of the proteorhodopsin appears, and/or altering the spectrophotometic properties of the proteorhodopsin as compared with that of the wild-type.

A conserved amino acid residue of proteorhodopsin is an amino acid that is found in the equivalent position of the 81 proteorhodopsins as depicted in FIGURE 3. A conserved amino acid residue, which alters the photochemical property of the proteorhodopsin when substituted with a different amino acid, is important for this invention. FIGURE 3 shows the alignment of amino acid sequences of 81 natural proteorhodopsin variants. Examples of conserved amino acid residues (H75, R94, D227 of BAC31A8) are shown in FIGURE 3. Such conserved amino acid residues can affect the conformation of the protein and the positioning of the all-trans-retinal molecule in relation to the proteorhodopsin protein.

The present invention identifies that the conserved histidine and arginine residues of a proteorhodopsin variant are important for the proton transport during a photocycle. The

conserved histidine residue (for example, H75 in Bac31A8 and H77 in Hot75m1) and the conserved arginine residue (for example, R94 in Bac31A8 and R96 in Hot75m1) are purportedly located near both the D97 residue and the all-trans-retinal molecule. These residues in proteorhodopsin interact with D97 or with another amino acid or a water molecule that interacts with D97, but do not interact directly with all trans-retinal. The conserved histidine and the conserved arginine are likely to be part of the hydrogen bondable groups responsible for the spectral change at different pH values.

Proteorhodopsin mutants, which have the conserved histidine substituted with an amino acid capable of forming a hydrogen bond, have an altered photochemical property that shifts the pK_{rh} to a lower pH (more acidic) value than the wild-type proteorhodopsin. For example, mutations at H75 in Bac31A8 proteorhodopsin and at H77 in Hot75m1 proteorhodopsin lower the pK_{rh} value of proteorhodopsin and expand the pH range of the basic form. In some H75 or H77 proteorhodopsin mutants, only the basic form exists in the pH range tested (such as pH 4.9-9.1). This is likely caused by the pK_{rh} values being shifted so low that it is outside the pH value used in the experiment. It is possible that an acid form will appear at a lower pH. Such mutants in the basic form are able to undergo a photocycle that results in a proton transport and go into the M-state with photo stimulation.

Amino acids capable of forming a hydrogen bond and suitable for substituting histidine in this invention include asparagine, glutamine, lysine, arginine, serine, theonine, tyrosine, aspartic acid (in its protonated state at acid pH), glutamic acid (in its protonated state at acidic pH), tryptophan, and any synthetic amino acid that has a functional group that is able to contribute a hydrogen to form a hydrogen bond. Preferred proteorhopsin mutants have the conserved histidine residue substituted with glutamine (Q), asparagine (N), glutamic acid (E), lysine (K), aspartic acid (D), and tryptophan (W).

Proteorhodopsin mutants, which have the conserved arginine substituted with a different amino acid, have an altered photochemical property that results in less difference in maximum absorption wavelength between a basic and an acidic form than the wild-type proteorhodopsin. Preferred proteorhopsin mutants have the conserved arginine residue mutated to alanine, glutamic acid or glutamine.

The altered photochemical property of proteorhodopsin mutants can be identified by measuring the spectral shift of the proteorhodopsin in intact or whole cells, wherein the proteorhodopsin is present in the cytoplasmic membrane, or in a solubilized polypeptide form stabilized by an amphipathic molecule, or in a pure polypeptide form. In addition, the altered

photochemical property can be identified by measuring the spectral shift of the proteorhodopsin in a membrane preparation, including but not limited to crude or partly purified membranes. Membrane preparations contain lipids and potentially membrane proteins other than proteorhodopsin.

The present invention provides a method for preparing a proteorhodopsin mutant having improved optical characteristics. The method comprises the steps of: (a) identifying a conserved amino acid residue of a wild-type proteorhodopsin variant, (b) mutagenizing the conserved amino acid residue and obtaining proteorhodopsin mutants, (c) determining the optical characteristics of the proteorhodopsin mutants, and (d) selecting the proteorhodopsin mutant having improved optical characteristics. The conserved amino acid residue, for example, is a histidine or an arginine residue. The wild-type proteorhodopsin variant can by mutagenized by any method, including but not limited to site-directed mutagenisis, known to a skilled person.

Nucleotide/Amino Acid Sequence of Proteorhodopsin Variants

The nucleotide and amino acid sequences of various proteorhodopsin variants (Figures 1-1 to 1-29) are deposited with Genbank under accession numbers AF349976-AF350003 and AF279106 (which contains the nucleotide sequence of a 105 kilobase genomic region that includes the gene encoding the Bac31A8 proteorhodopsin). The nucleotide and amino acid sequences of newly isolated proteorhodopsin variants (Figures 1-30 to 1-51) are deposited with Genbank under accession numbers AY210898-AY210919. The nucleotide and amino acid sequences of newly isolated proteorhodopsin variants (Figures 1-52 to 1-79) are deposited with Genbank under accession numbers AY250714-AY250741. The nucleotide and amino acid sequences of newly isolated proteorhodopsin variants (Figures 1-80 to 1-81) are deposited with Genbank under accession numbers AY372453 and AY372455. All these sequences are specifically incorporated herein by reference in this application. Any of these natural proteorhodopsin amino acid sequences can be used to make the proteorhodopsin mutants. For example, the natural proteorhodopsin variant include Hot75m1, Bac31A8, Bac40E8, Bac41B4, Bac64A5, Hot0m1, Hot75m3, Hot75m4, Hot75m8, MB0m1, MB0m2, MB20m2, MB20m5, MB20m12, MB40m1, MB40m5, MB100m5, MB100m7, MB100m9, MB100m10, PalB1, PalB2, PalB5, PalB7, PalB6, PalB8, PalE1, PalE6, PalE7, MED 26, MED27, MED36, MED101, MED102, MED106, MED25, MED202, MED204. MED208, REDA9, REDB9, REDF9, RED19, RED2, RED23, RED27, RED30, RED4, RED5,

REDr6a5a14, REDr6a5a6, REDr7_1_4, REDs3_7, REDr7_1_15, REDs3_15, medA15r8ex6, REDr7_1_16, medA15r11b9, medA15r9b5, medA15r8b3, medA15r11b3, medA15_r8_1, medA17R9_1, medA15r8b9, medA19_R8_16, medA19_R8_19, medA17_R8_6, medA15r9b7, medA15_R8_3, medA15r10b5, medA19_r9_9, medA15_r8ex7, medA19_R8_20, medA15_R8ex9, medA15_r9_3, medA17_r8_15, medA17_r8_11, medA15r8b8, medA15r8ex4, ANT32C12 PR and HOT2C01 PR. Preferred natural proteorhodopsin variants are Hot75m1 and Bac31A8 proteorhodopsin.

The nucleotide sequence of any of the above proteorhodopsin genes can be obtained or derived using the method of Béjà, *et al.* (2000) or WO 01/83701. An example of a gene encoding a proteorhodopsin variant is the bacterioplankton Bacterial Artificial Chromosome (BAC) clone Bac31A8 (also known as EBAC31A08) (see WO 01/83701). One skilled in the art can obtain other genes encoding different proteorhodopsin variants using the identical techniques described above. One skilled in the art can also obtain or clone proteorhodopsin genes from organisms obtained from natural habitats by designing primers and using degenerate PCR, heterologous hybridization, or random sequencing of DNA.

Suitable Host

The proteorhodopsin is capable of proper folding and integrating into a membrane when expressed or synthesized in a suitable host. A suitable host is a cell that naturally expresses proteorhodopsin, such as the SAR 86 strain. A suitable host can also be a cell that naturally does not express proteorhodopsin. A suitable host includes a marine and a non-marine bacteria. Preferably, a suitable host is deficient in any outer-membrane protease, either naturally or constructed not to express such protease. An example of a suitable host is an eubacteria cell; preferably, a Gram-negative bacteria; for example, a proteobacteria such as a gamma-proteobacteria. The gamma-proteobacteria can belong to the Enterobacteriaceae family; such as *Escherichia*, *Edwardsiella*, *Citrobactor*, *Salmonella*, *Shigella*, *Klebsiella*, *Enterobacter*, *Serratia*, *Proteus*, and *Yersinia*. In the embodiment, the gamma-proteobacteria belonging to the *Salmonella* genus is *Salmonella typhimurium*. In another embodiment, the gamma-proteobacteria belonging to the *Escherichia* genus is *Escherichia coli*.

Optionally, the polypeptide is expressed or synthesized in a strain that is an outer membrane protease-deficient strain. Such suitable strains include the *E. coli* strains: UT5600 (as disclosed by Béjà, *et al.* (2000), and Dioumaev, *et al.*) and BL21-Codonplus-RIL. Alternatively, the gamma-proteobacteria is the SAR86 strain.

The present invention also provides for a host cell comprising a polynucleotide encoding the proteorhodopsin mutant. The suitable host cell is capable of expressing or synthesizing a functional proteorhodopsin from the polynucleotide. A functional proteorhodopsin is a polypeptide capable of undergoing a photocycle in a suitable environment.

The one or more mutations can be constructed by site-directed mutagenesis of a cloned proteorhodopsin gene. Specific designed mutations can be constructed or variability can be introduced to produce a variety of mutants. Such mutagenesis and other methods for manipulation of the nucleic acid and protein are well known to one skilled in the art and are described by Sambrook, *et al.* (*Molecular cloning: a laboratory manual.* 3d ed. Cold Spring Harbor Laboratory Press, Cold Spring harbor, NY).

Suitable Vector

The present invention provides a polynucleotide encoding the proteorhodopsin mutant. One skilled in the art is able to construct a nucleotide sequence using the information of the amino acid sequence of the polypeptide, based on the universal genetic code. The polynucleotide optionally comprises a promoter operatively linked 5' to the open reading frame encoding the polynucleotide. The polynucleotide further comprises appropriate promoter control, translational control and stop and other necessary sequences in order to express that polypeptide in a suitable host.

The polynucleotide can be a chromosome, an episome, a plasmid, or any suitable expression vector. The polynucleotide is capable of amplification in a host cell. For an eubacterial host cell, a wide variety of expression vectors can be used for introduction by transformation, conjugation, transduction or transfection of the polynucleotide into the eubactaerial host cell. Vectors include plasmids, such as pBR322, pMB9, pBAD-TOPO® (from the pBAD TOPO® TA Expression Kit, Invitrogen, La Jolla, CA), pTrcHis2-TOPO® (from the pTrcHis2-TOPO® TA Expression Kit, Invitrogen, La Jolla, CA) and the like; cosmids, such as pVK100, and the like; and viruses such as P22, and the like.

Isolation of the Proteorhodopsin from a Host

The proteorhodopsin polypeptide can be purified from a host cell using a variety of methods, which follow a general scheme:

1. Lyse the host cells containing the polypeptide.

2. Dissolve the membrane with detergent.

3. Load impure polypeptide on His-tag affinity resin.

4. Wash resin to remove impurities.

5. Remove protein from column.

6. Exchange buffer and concentrate protein.

Methods of purification can differ for each intended use of the protein since requirements of yield and purity differ in each application.

Lysis can be carried out by sonication, osmotic shock, freeze thaw or French press or a combination of above methods. A preferred method of lysis is freeze-thaw followed by French press. Insoluble materials are removed by centrifugation (e.g. 1500 xg, 30 min.) and the supernatant collected. The supernatant is further centrifuged (e.g. 150,000 xg. 1 hour) and the pellet containing primarily membrane material is collected.

The membrane material can be solubilized with detergent. Useful detergents include, but are not limited to, dodecyl-β-maltoside, octyl-β-glucoside, or Triton® X-100. Dodecyl-β-maltoside (2%) is a preferred detergent.

The solubilized membrane material is incubated with His-tag affinity resin. For example, the solubilized membrane material is incubated with Talon™ resin from Clontech (Palo Alto, CA) overnight at 4 °C with gentle agitation.

The solubilized membrane material is purified from contaminating proteins by washing the resin, for example, with three column volumes of buffer containing 0.1% dodecyl-β-maltoside three times.

The purified polypeptide can be removed from His-tag purification resins by a variety of methods including, but not limited to, incubation with EDTA, incubation in acidic pH, or incubation at high temperature. A preferred method is to incubate the resin with a buffer containing 0.5 M EDTA.

The buffer can be exchanged by a variety of methods including, but not limited to, concentration and dilution using membrane based concentrators, dialysis, or desalting columns. In a preferred method, the protein is concentrated using a membrane based centrifugation concentrator device and then diluted into a desired buffer. The process is repeated to completely exchange the buffer. Then the protein is concentrated again using the same concentrator device.

Further methods for the isolation of the polypeptide in the crude or partly pure form are disclosed by Dioumaev, *et al.* and Krebs, *et al.*, both of which are incorporated in this application by reference.

Specific Embodiments of the Invention

In one embodiment of the present invention, the proteorhodopsin mutant is derived from Hot75m1 proteorhodopsin variant with the conserved histidine substituted with a glutamine residue. The polypeptide is expressed from an *E. coli* host cell. When analyzed for its absorbance behavior in the membrane of intact cells, the mutant polypeptide exhibits a spectral shift with the pK_{rh} value of less than about 5.8 (as compared with 7.2 of the wild-type proteorhodopsin). The shift of pK_{rh} is about 1.4 units.

In another embodiment of the present invention, the proteorhodopsin mutant is derived from Hot75m1 proteorhodopsin variant with the conserved histidine substituted with an asparagine residue. The polypeptide is expressed from an *E. coli* host cell. When analyzed for its absorbance behavior in the membrane of intact cells, the polypeptide exhibits a spectral shift with the pK_{rh} value of less than about 5.1. The shift in the pK_{rh} is about 2.1 units.

In another embodiment of the present invention, the proteorhodopsin mutant is derived from Hot75m1 proteorhodopsin variant with the conserved histidine substituted with a lysine residue. The polypeptide is expressed from an *E. coli* host cell. When analyzed for its absorbance behavior in the membrane of intact cells, the polypeptide does not exhibit a spectral shift (or does not titrate) when analyzed at a pH range from 4.9 to 9.1. The shape of the curves obtained for this mutant proteorhodopsin, when absorbance of the polypeptide is plotted against the wavelength of the light source, is similar to the shape for the curve obtained from the basic form of the wild-type Hot75m1 proteorhodopsin. Also, the wavelength of the maxima absorbance of this mutant proteorhodopsin (about 495 nm) is much closer to the maxima of absorbance for the basic form of Hot75m1 (about 490 nm) than the acidic form of Hot75m1 (about 540 nm). Based on this reasoning, this mutant proteorhodopsin is likely in its basic form when the pH of its environment is from 4.9 to 9.1. It is likely that the pK_{rh} of this mutant is shifted to a pH much lower than pH 4.9.

Technical Application of the Subject Polypeptide

The proteorhodopsin mutants taught herewith have many technical applications. For example, the proteohodopsin mutant can be incorporated into instruments or devices having photochromic applications, photoelectric applications, and/or phototransport applications.

Under photochromic applications, the polypeptide can be used for its light absorption properties for optical data storage, interferometry and/or photonics. Photochromic applications include, but are not limited to, holographic film. The proteorhodopsin mutant can be used as a photochromic all-trans-retinal protein for optical data storage devices. The proteorhodopsin mutant can be used in a device for information storage, such as 2-D storage, 3-D storage, holographic storage, associative storage, or the like. The proteorhodopsin mutant can be used in a device for information processing, such as optical bistability/light switching, optical filtering, signal conditioning, neural networks, spatial light modulators, phaseconjugation, pattern recognition, interferometry, or the like.

The present invention also provides a method of storing and retrieving optical data using the proteorhodopsin mutant of the invention. A method of storing and retrieving optical data, for example, comprises the steps of: (a) providing a film comprising the proteorhodopsin mutant immobilized in a matrix; (b) recording the optical data by exposing the film to light of a wavelength that is absorbed by the proteorhodopsin mutant in a predetermined pattern and selectively converting portions of the film to an excited state; (c) exposing the film of step (b) to light of a wavelength that is absorbed either by the basal or the excited state of the proteorhodopsin mutant; (d) detecting the optical data by an optical recording device. The detecting of step (d) can be conducted by any optical recording device. The optical recording device can be a video camera, which can be a charged coupled device (CCD). This method as applied to bacteriorhodopsin, and mutants thereof, is known in the art (see patents listed in Table 1) and can be readily modified by replacing the bacteriorhodopsin with the proteorhodopsin mutant of the invention.

Under photoelectric applications, the proteorhodopsin mutant can be used in devices for its light-induced changes in electric fields caused by proton transport, such as in ultrafast light detection, artifical retina, and/or light/motion detectors.

Under phototransport applications, the proteorhodopsin mutant can be used for its light-induced proton transport across a membrane, such as photovoltaic device. One such photovoltaic device is a light-driven energy generator comprising the polypeptide, whereby light energy can be converted to chemical energy. The manufacture and use of such a light-

driven energy generator is disclosed in WO 01/83701. The proteorhodopsin mutant can also be used in devices for ATP generation in reactors, desalination of seawater, and/or conversion of sunlight into electricity.

The present invention provides for a light-driven energy generator comprising the use of the proteorhodopsin mutant. A light-driven energy generator comprises: (a) the proteorhodopsin mutant of the present invention, (b) a cell membrane, (c) a source of all-trans-retinal, and (d) a light source, wherein the proteorhodopsin mutant integrates within the cell membrane to produce an integrated proteorhodopsin mutant, and the integrated proteorhodopsin mutant binds covalently to all-trans-retinal to produce a light absorbing pigment. Light-driven energy generators as applied to natural proteorhodopsin variants are disclosed by WO 01/83701, and can be modified by replacing the proteorhodopsin with the proteorhodopsin mutants. The light-driven generator utilizes the proteorhodopsin mutant to convert light-energy into biochemical energy. The light-driven energy generator takes advantages of the functional properties of a proteorhodopsin mutant, which is expressed and correctly folded in a suitable host cell.

The proteorhodopsin mutant can also be used in devices for 2D harmonic generation, radiation detection, biosensor applications, or the like.

Hampp (*Chem. Rev.*,100: 1755-76 (2000)) discloses various uses and instruments and devices that utilize the photochemical properties of bacteriorhodopsin. Proteorhodopsin mutants of this invention can be used in the analogous manner as described by Hampp, which is specifically incorporated by reference.

Further, proteorhodopsin mutants can be used to replace bacteriorhodopsin (BR) for a variety of devices/processes that utilize bacteriorhodopsin, for example, in the following list of patents (Table 1), which are incorporated herein by reference.

TABLE 1.

Inventor(s)	Priority	Patent No(s).	Subject
Inatomi and Isoda	26 APR 1983	JP 59197849	Photosensor
Isoda	02 MAR 1984	JP 60185228	Recording and reproducing method of information
Isoda and Daimon	02 MAR 1984	JP 60184246	Light recording medium
Inatomi	19 NOV 1984	JP 61124384	ATP reproduction bioreactor
Hikima, *et al.*	09 JUL 1985	JP 62011158	Biochemical element

Inventor(s)	Priority	Patent No(s).	Subject
Sora, *et al.*	14 SEP 1985	JP 62063823	Optical sensor
Arai, *et al.*	08 OCT 1986	JP 63092946 US 4965174	Recording medium and process for forming color image with same
Arai, *et al.*	08 OCT 1986	JP 63092947	Protein-enzyme biochemical optical recording medium and imaging process
Inoue	20 MAR 1987	JP 63231424	Optical switch formed by using BR
Yamamoto, *et al.*	20 MAR 1987	JP 63231337	Recording medium and image forming method using same
Ogawa	29 MAY 1987	JP 63299374 US 4896049	Color image sensor obtained from visual photosensitive material; derived from biological substances
Inoue	20 JUL 1987	JP 1023125	Image pickup element
Oesterhelt, *et al.*	10 SEP 1987	EP 306985	BR modifications and methods for their preparation
Iwashita, *et al.*	29 OCT 1987	JP 1116536	Recording medium and color image forming device using same
Kawada, *et al.*	21 DEC 1987	JP 1165186	Switching device
Oyama, *et al.*	28 MAR 1988	JP 1245810	Ion permeating membrane and its control by light irradiation
Sakakibara and Fukuda	12 OCT 1988	JP 2104600	Photo-sensitive pigment, production thereof, optical recording material and colorant
Oyama, *et al.*	18 MAR 1989	JP 2247233	Ion transmission membrane and ion transfer method using the membrane
Ikematsu	20 MAR 1989	JP 2247894	Plastic optical memory element
Tokunaga and Nakasako	27 MAR 1989	JP 2251949	Photoreactive composition
Watanabe	26 MAY 1989	JP 2310538	Optical switch using BR
Watanabe	24 JUL 1989	JP 3054532	Optical function element
Koyama	25 AUG 1989	JP 3081756	Optical recording material
Takeda, *et al.*	28 AUG 1989	EP 417541	Process for preparing a rhodopsin-coated membrane for use in transducers
Miyasaka and Shizukuishi	13 SEP 1989	JP 3100524 US 5279932	Optical response element
Miyasaka	18 OCT 1989	JP 3205520 US 5107104	Photoelectric transducer having photosensitive chromoprotein film, i. e. BR

Inventor(s)	Priority	Patent No(s).	Subject
Koyama and Yamaguchi	29 NOV 1989	JP 3170500	Orientation of photosensitive pigment protein
Miyasaka	18 DEC 1989	JP 3188421	Optical information-converting element
Koyama, *et al*	02 MAR 1990	JP 3252530	Color image photodetector
Hampp, *et al.*	08 MAR 1990	EP 445758 US 5223355	Methods for improving the signal-to-noise ratio in holography when using br-based recording media
Oyama, *et al.*	13 APR 1990	JP 3295278	Photoelectric conversion element
Miyasaka	25 APR 1990	JP 4006420	Photoelectric conversion element
Miyasaka and Kitaguchi	25 APR 1990	JP 4009400	Fixation of photosensitive chromoprotein
Saito, *et al.*	08 JUN 1990	JP 4042585	Photoresponsive excitable synthetic membrane and its manufacture
Saito, *et al.*	31 JUL 1990	JP 4088995	ATP synthesizing device
Miyasaki and Koyama	28 NOV 1990	JP 3237769	Color picture image sensor
Birge and Lawrence	23 JAN 1991	US 5228001	Optical random access memory
Saito, *et al.*	18 FEB 1991	JP 4262583	Photoresponsive excitability artificial membrane and manufacture thereof
Miyasaka	11 APR 1991	JP 4312078 JP 4312079 JP 4312080 JP 4312081 US 5260559	Image information detecting method by photoelectric conversion element
Birge and Govender	20 APR 1991	US 5253198	Three-dimensional optical memory
Saito, *et al.*	21 AUG 1991	JP 05048176	Optical responsive exciting artificial film and method for manufacturing thereof
Saito, *et al.*	12 NOV 1991	JP 5133795	Bio-element
Fukuzawa and Kuwano	12 NOV 1991	JP 5130880	Hydrogen generator
Fukuzawa	14 NOV 1991	JP 5227765	Micro electrostatic actuator
Fuktuawa and Kuwano	14 NOV 1991	JP 5136483	Image detection method and image sensors
Fukuzawa and Kuwano	14 NOV 1991	JP 5134187	Close-view microscope

Inventor(s)	Priority	Patent No(s).	Subject
Fukuzawa and Kuwano	14 NOV 1991	JP 5134188	Close-view microscope
Haronian and Lewis	07 FEB 1992	US 5248899	Neural network using photoelectric substance for storing or retrieving information
Miyasaka and Yamaguchi	27 JAN 1992	JP 5204090	Production of BR-oriented film
Miyasaka	22 MAY 1992	JP 5322645	Light receiving element
Lewis, *et al.*	30 NOV 1992	US 5346789 WO 9311470	Oriented biological material for optical information storage and processing
Miyasaka	09 DEC 19	92 JP 6174544	Photodetector
Fukuzawa	17 DEC 1992	JP 6235606	Position detector
Miyasaka	21 DEC 1992	JP 6186078	Photoelectric conversion material
Takei and Shimzu	24 DEC 1992	JP 6194612 US 5618654	Photo-controlled spatial light modulator
Tomita, *et al.*	09 FEB 1993	JP 6234626	Polymerizable proteoliposomes
Koyama	12 APR 1993	JP 6294682	Photoelectric conversion element
Takei and Shimizu	10 JUN 1993	JP 6347835	Production of violet thin film
Fukuzawa, *et al.*	04 OCT 1993	JP 7106616	Optical energy conversion thin film and production thereof
Kato and Tanaka	09 MAY 1994	JP 7301591	Moisture measuring method
Dyukova and vsevolodov	26 MAY 1994	US 5518858	Photochromic compositions and materials containing BR
Lewis, *et al.*	09 JUN 1994	US 5470690	Optical information storage on a BR-containing film
Ikematsu and izeki	16 NOV 1994	JP 8146469	Optical information conversion element
Birge	27 DEC 1994	US 5559732 WO 9621228	Branched photocycle optical memory device
Fitzpatrick	23 FEB 1995	US 5563704	Camera and method for holographic interferometry using an erasable photosensitive photopolymer film
Otomo	22 MAR 1995	JP 8261980	Ion sensor
Kolodner and Rousseau	05 APR 1996	US 5781330	High efficiency optical switching and display devices
Rao, *et al.*	05 JUN 1996	US 5757525 WO 9746907	All-optical devices
Rao, *et al.*	09 SEP 1996	US 5854710 WO 9810315	Optical fourier processing

Inventor(s)	Priority	Patent No(s).	Subject
Fiedler, *et al.*	25 OCT 1996	WO 9819217	Method to prepare the production of structured metalcoatings using proteins
Rayfield and Hsu	06 MAY 1997	US 5825725	Method and apparatus for reversible optical data storage

The following examples further illustrate the present invention. These examples are intended merely to be illustrative of the present invention and are not to be construed as being limiting.

EXAMPLES

EXAMPLE 1. Cloning of Proteorhodopsins Into Expression Vector.

A total of 21 of the 81 known natural proteorhodopsin genes were cloned in the pTrcHis2 expression vector. Different natural proteorhodopsin genes were PCR amplified using Taq DNA Polymerase (Roche Applied Science) as described by the manufacturer. The primers used to amplify all proteorhodopsin variants were PR-u4 and PR-d2 (see Table 2) and the templates were pCR2.1 containing the relevant proteorhodopsin (Béjà, *et al.*, 2000; Béjà, *et al.*, 2001). The PCR products were cloned in the pTrcHis2 vector using the pTrcHis2 TOPO TA Expression Kit (Invitrogen, Carlsbad, CA), as described by the manufacturer. Restriction enzyme digestions of the plasmids were used to identify clones containing the insert in the correct orientation. The plasmids were sequenced using the primers pTrcHis Forward and pTrcHis Reverse to ensure that no sequence variations were introduced during the PCR and cloning procedure. The plasmids contain the pTrc promoter transcribing the proteorhodopsin gene with a C-terminal extension containing a *myc* epitope and six histidine residues. A map of one of the expression plasmids is shown in FIGURE 4.

TABLE 2. Sequences of Oligonucleotide Primers Used

Primer name	Sequence (5' to 3')
PR-u4 (SEQ. ID NO : 185)	AAATTATTACTGATATTAGGTAGTG
PR-d2 (SEQ ID NO : 186)	AGCATTAGAAGATTCTTTAACAGC
pTrcHis	GAGGTATATATTAATGTATCG

Primer name	Sequence (5' to 3')
Forward (SEQ ID NO : 187)	
pTrcHis Reverse (SEQ ID NO : 188)	GATTTAATCTGTATCAGG
31A8-m6A (SEQ ID NO : 189)	TGTTACTGGTATTGCTTTCTGG**aATT**ACATGTACATGAGAGGGGT
31A8-m6B (SEQ ID NO : 190)	ACCCCTCTCATGTACATGTA**ATt**CCAGAAAGCAATACCAGTAACA
31A8-m7A (SEQ ID NO : 191)	TGTTACTGGTATTGCTTTCTGG**cAg**TACATGTACATGAGAGGGGT
31A8-m7B (SEQ ID NO : 192)	ACCCCTCTCATGTACATGTA**cTg**CCAGAAAGCAATACCAGTAACA
31A8-m8A (SEQ ID NO : 193)	TGTTACTGGTATTGCTTTCTGG**aAa**TACATGTACATGAGAGGGGT
31A8-m8B (SEQ ID NO : 194)	ACCCCTCTCATGTACATGTA**tTt**CCAGAAAGCAATACCAGTAAC
H75m1-m2A (SEQ ID NO : 195)	TTACTGGTATAGCTTTTTGG**aAT**TATCTCTATATGAGAGGTGTTT
H75m1-m2B (SEQ ID NO : 196)	AAACACCTCTCATATAGAGATA**ATt**CCAAAAAGCTATACCAGTAA
H75m1-m3A (SEQ ID NO : 197)	TTACTGGTATAGCTTTTTGG**CAg**TATCTCTATATGAGAGGTGTTT
H75m1-m3B (SEQ ID NO : 198)	AAACACCTCTCATATAGAGATA**cTG**CCAAAAAGCTATACCAGTA A
H75m1-m4A (SEQ ID NO : 199)	TTACTGGTATAGCTTTTTGG**aAa**TATCTCTATATGAGAGGTGTTT
H75m1-m4B (SEQ ID NO : 200)	AAACACCTCTCATATAGAGATA**tTt**CCAAAAAGCTATACCAGTAA
H75m1-m7A (SEQ ID NO : 201)	CTGGTGATACCCCAACAGTATTC**gcA**TATATTGATTGGTTATTAA

Primer name	Sequence (5' to 3')
H75m1-m7B (SEQ ID NO : 202)	TTAATAACCAATCAATATA**Tgc**GAATACTGTTGGGGTATCACCAG
H75m1-m8A (SEQ ID NO : 203)	CTGGTGATACCCCAACAGTATTC**caA**TATATTGATTGGTTATTAA
H75m1-m8B (SEQ ID NO : 204)	TTAATAACCAATCAATATA**Ttg**GAATACTGTTGGGGTATCACCAG
H75m1-m9A (SEQ ID NO : 205)	CTGGTGATACCCCAACAGTATTC**gaA**TATATTGATTGGTTATTAA
H75m1-m9B (SEQ ID NO : 206)	TTAATAACCAATCAATATA**Ttc**GAATACTGTTGGGGTATCACCAG
H75m1-m12A (SEQ ID NO : 207)	TTACTGGTATAGCTTTTTGG**gAT**TATCTCTATATGAGAGGTGTTT
H75m1-m12B (SEQ ID NO : 208)	AAACACCTCTCATATAGAGATA**ATc**CCAAAAAGCTATACCAGTA A
H75m1-m13A (SEQ ID NO : 209)	TTACTGGTATAGCTTTTTGG**gaa**TATCTCTATATGAGAGGTGTTT
H75m1-m13B (SEQ ID NO : 210)	AAACACCTCTCATATAGAGATA**ttc**CCAAAAAGCTATACCAGTAA
H75m1-m15A (SEQ ID NO : 211)	TTACTGGTATAGCTTTTTGG**tgg**TATCTCTATATGAGAGGTGTTT
H75m1-m15B (SEQ ID NO : 212)	AAACACCTCTCATATAGAGATA**cca**CCAAAAAGCTATACCAGTAA

The pTrcHis Forward and pTrcHis Reverse oligonucleotides were obtained from Invitrogen, the rest of the oligonucleotides were purchased from Operon, the primers used for site-directed mutagenesis were PAGE purified. Bold letters indicate the histidine or arginine codons that were mutated, lowercase letters indicate nucleotides that were changed compared to the wild-type sequence.

Measurement of Proteorhodopsin Spectra in Intact Cells

The proteorhodopsin expression plasmids and the control plasmid pTrcHis2-LacZ (Invitrogen) were transformed into competent cells of the strain BL21-Codonplus-RIL

(Stratagene) as described by the manufacturer. The transformed cells were plated on LA + 0.5% glucose + 100 μg/ml carbenicillin + 25 μg/ml chloramphenicol plates and incubated overnight at 37°C. Cells from these plates were grown in 6 ml LB + 0.5% glucose + 100 μg/ml carbenicillin + 25 μg/ml chloramphenicol + 10 μM all-trans-retinal medium in glass tubes at at 37°C for 6 hours, where the cells reached early stationary phase. The cells constitutively express proteorhodopsins from the uninduced pTrc promoter. The cells were collected by centrifugation at 4,500xg for 6 minutes, the medium was discarded and the cells were resuspended in 5 ml sterile water. A 900 μl sample of cells were added to a 100 μl aliquot of concentrated buffer (1.0 M acetate (pH 4.9), 1.0 M MES (pH 5.8), 1.0 M MOPS (pH 6.7), 1.0 M TAPS (pH 8.1), or 1.0 M CHES (pH 9.1)). The spectra of the 21 natural proteorhodopsins were measured in intact cells at five different extracellular pH values using a spectrophotometer (On-Line Instrument Systems, Inc.) adapted for use with turbid samples by placing the photomultiplier detector adjacent to the sample cuvette. The absorption from a strain containing a control plasmid (pTrcHis2-LacZ) was subtracted from proteorhodopsin containing samples. To correct for differences in background light scattering caused by variations in cell densities between the sample and the reference, a linear baseline determined by least-squares fitting of the first ten and last ten wavelength and absorbance values in each spectrum was subtracted from each spectrum. The spectra were then adjusted to the same minimum value by subtracting the differences in minimum values.

Typical proteorhodopsin spectra, such as Bac31A8 and Hot75m1 inside intact cells at different pH values, can be constructed by plotting absorbance vs. wavelength. Two different spectral forms of the proteorhodopsin variants (Bac31A8 and Hot75m1) are observed under acidic conditions (pH 4.9) and basic conditions (pH 9.1). At intermediate pH values, spectra representing a mixture of the acidic and basic forms are observed.

To determine the pK_{rh} value, the different adjusted absorbance intensities at each wavelength and pH were fitted to equation 1 by least-squares regression analysis using the Solver function of Microsoft Excel.

$$Abs = Abs_{Acidic} + Abs_{Basic-Acidic} * (10^{pH-pK_{rh}}) \quad (1)$$

A different value of Abs_{Acidic} and $Abs_{Basic\text{-}Acidic}$ were fitted at each wavelength, but the same value of pK_{rh} was fitted for data at all wavelengths. The values Abs_{Acidic} and $Abs_{Basic\text{-}}$

$_{Acidic}$ were used to reconstruct wavelength spectra for Acidic and Basic forms of proteorhodopsin and from this data the wavelength maximum at acidic and basic pH were determined (Table 3). The wild-type Bac31A8 proteorhodopsin in intact cells had absorbance maximum at 538 nm at acidic conditions and at 519 nm at basic conditions. The pK_{rh} value of the titratable group(s) responsible for the spectral shift was determined to be 7.6. The wild-type Hot75m1 proteorhodopsin had absorbance maximum at 546 nm at acidic conditions and at 493 nm at basic conditions. The value of the pK_{rh} of the titratable group(s) responsible for the spectral shift was determined to be 8.2.

Table 3 summarizes the results obtained with the 21 different proteorhodopsin genes expressed and characterized:

TABLE 3. Spectral Property of Proteorhodopsin Variants

Protein	Name of expression plasmid	pK_{rh}	Basic Max (nm)	Acidic Max (nm)
Bac31A8	pTrcHis2-Bac31A8	7.6	521	538
Bac40E8	pTrcHis2-Bac40E8	7.7	519	538
Bac64A5	pTrcHis2-Bac64A5	7.6	519	538
Hot0m1	pTrcHis2-Hot0m1	8.0	518	538
Hot75m1	pTrcHis2-Hot75m1	8.2	493	546
Hot75m3	pTrcHis2-Hot75m3	7.2	488	538
Hot75m4	pTrcHis2-Hot75m4	7.6	490	538
Hot75m8	pTrcHis2-Hot75m8	8.0	493	538
MB0m1	pTrcHis2-MB0m1	7.9	518	540
MB0m2	pTrcHis2-MB0m2	7.9	523	540
MB20m2	pTrcHis2-MB20m2	7.9	523	538
MB20m5	pTrcHis2-MB20m5	8.2	526	569
MB20m12	pTrcHis2-MB20m12	7.6	524	540
MB40m1	pTrcHis2-MB40m1	7.7	519	538
MB40m5	pTrcHis2-MB40m5	8.5	525	558
MB40m12	pTrcHis2-MB40m12	7.5	523	536
MB100m5	pTrcHis2-MB100m5	7.8	523	538

MB100m7	pTrcHis2-MB100m7	8.1	524	550
MB100m9	pTrcHis2-MB100m9	7.3	524	538
MB100m10	pTrcHis2-MB100m10	7.7	524	538
PalE6	pTrcHis2-PalE6	7.1	490	542

We have expressed and measured the spectral properties at different pH values for 21 of the 81 known proteorhodopsins. All of them have different absorption spectra at acidic and basic conditions and the pK_{rh} value is in the pH 7-8.5 range for all of the natural proteorhodopsin tested. Thus, the pH dependent shift in spectral properties is a property of the entire proteorhodopsin family.

EXAMPLE 2. Mutagenesis of Proteorhodopsins

To test whether a conserved histidine residue (H75 in Bac31A8 and H77 in Hot75m1) and a conserved arginine residue (R96 in Hot75m1) are involved in forming the protonatable group responsible for the pH dependent change in the spectral properties of proteorhodopsin variants, these amino acid residues in Bac31A8 and Hot75m1 were mutagenized using the QuickChange Site-Directed Mutagenesis Kit (Stratagene) as described by the manufacturer

Table 4 summarizes the templates and primers used to construct the different mutants.

TABLE 4. Templates and Primers Used to Construct Mutants

Mutant	Name of expression plasmid	Template used	Primers used
Bac31A8 H75N	pTrcHis2-Bac31A8 H75N	pTrcHis2-Bac31A8	31A8-m6A and 31A8-m6B
Bac31A8 H75Q	pTrcHis2-Bac31A8 H75Q	pTrcHis2-Bac31A8	31A8-m7A and 31A8-m7B
Bac31A8 H75K	pTrcHis2-Bac31A8 H75K	pTrcHis2-Bac31A8	31A8-m8A and 31A8-m8B
Hot75M1 H77N	pTrcHis2-Hot75M1 H77N	pTrcHis2-Hot75M1	H75m1-m2A and H75m1-m2B
Hot75M1 H77Q	pTrcHis2-Hot75M1 H77Q	pTrcHis2-Hot75M1	H75m1-m3A and H75m1-m3B
Hot75M1 H77K	pTrcHis2-Hot75M1 H77K	pTrcHis2-Hot75M1	H75m1-m4A and H75m1-m4B
Hot75M1 R96A	pTrcHis2-Hot75m1 R96A	pTrcHis2-Hot75M1	H75m1-m7A and H75m1-m7B
Hot75M1 R96Q	pTrcHis2-Hot75m1 R96Q	pTrcHis2-Hot75M1	H75m1-m8A and H75m1-m8B
Hot75M1 R96E	pTrcHis2-Hot75m1 R96E	pTrcHis2-Hot75M1	H75m1-m9A and H75m1-m9B
Hot75M1 H77D	pTrcHis2-Hot75M1 H77D	pTrcHis2-Hot75M1	H75m1-m12A and H75m1-m12B
Hot75M1 H77E	pTrcHis2-Hot75M1 H77E	pTrcHis2-Hot75M1	H75m1-m13A and H75m1-m13B
Hot75M1 H77W	pTrcHis2-Hot75M1 H77W	pTrcHis2-Hot75M1	H75m1-m15A and H75m1-m15B

The plasmids were sequenced using the pTrcHis Forward primer (Invitrogen) to ensure that the desired mutations had been introduced into the plasmids.

These proteorhodopsin mutants were expressed; their spectra in intact cells at different pH values were measured and the pK_{rh} values were calculated as described above.

(a) Histidine Residue

Table 5 summarizes the pK_{rh} values and the absorption maxima of the acidic and basic spectra of the wild-type and mutants of Bac31A8.

TABLE 5. pK_{rh} and Absorption Maxima of BAC31A8 Mutants

Protein	pK_{rh}	Basic Max (nm)	Acidic Max (nm)
Bac31A8	7.6	519	538
Bac31A8 H75Q	6.0	518	540
Bac31A8 H75N	NA	521	-
Bac31A8 H75K	NA	522	-

Note: "NA" means that no significant pH titration is observed, so a pK_{rh} value cannot be determined. "–" means that only a single absorption maximum can be reported because of the lack of titration in the pH range tested.

These results show that asparagine or glutamine mutations in the conserved histidine residue of Bac31A8 resulted in a value of the pK_{rh} that was shifted towards acidic pH. The Bac31A8 H75Q mutant has a pK_{rh} value of 6.0 compared with 7.6 observed with the wild-type protein. Both the Bac31A8 H75N and Bac31A8 H75K mutants show no or very little pH dependent change in spectral properties, the mutant proteins appear to be only in the basic form in the pH range tested here. It is likely caused by the pK_{rh} value being shifted so much to be outside the pH range of 4.9-9.1. The pK_{rh} value of the Bac31A8 H75N mutant could not be determined accurately from these data, but is estimated to be 5.8 or lower.

The spectra of Hot75m1 H77Q, Hot75m1 H77N, Hot75m1 H77K, Hot75m1 H77D, Hot75m1 H77E and Hot75m1 H77W at different pH values were determined. FIGURE 5

shows the pH-dependent change in absorption spectra of the acid and the basic form of wild-type Hot75m1 and mutant Hot75m1 H77Q. The pK_{rh} of the wild type is 8.2, which is lowered to 5.8 in the mutant. FIGURE 6 shows the fraction of the acid form of Hot75ml and Hot75ml H77Q at different pH. The basic form clearly predominates at pH 9 and above for the wild-type Hot75ml protein and at pH 7 and above for the Hot75ml H77Q protein.

Table 6 summarizes the pK_{rh} value and the maxima of the acidic and basic absorption spectra of the wild-type and six mutants.

TABLE 6. pK_{rh} and Absorption Maxima of Hot75M1 Mutants

Protein	pK_{rh}	Basic Max (nm)	Acidic Max (nm)
Hot75M1	8.2	493	546
Hot75M1 H77Q	5.8	487	516
Hot75M1 H77N	5.5	487	497
Hot75M1 H77K	NA	500	-
Hot75M1 H77D	NA	515	-
Hot75M1 H77E	5.3	487	529
Hot75M1 H77W	5.8	486	529

Note: "NA" means that no significant pH titration is observed, so a pK_{rh} value cannot be determined. "–" means that only a single absorption maximum can be reported because of the lack of titration in the pH range tested.

These results show that mutation in the conserved H77 histidine residue of Hot75m1 that changes it to asparagine, glutamine, glutamic acid or tryptophan resulted in a pK_{rh} value that was significantly shifted towards acidic pH compared with that of the wildtype Hot75m1 protein. The Hot75m1 H77Q and the Hot75m1 H77W mutants have a pK_{rh} value of 5.8 compared with 8.2 of the wild-type protein. The pK_{rh} value of the Hot75m1 H77N and the Hot75m1 H77E mutants were even lower, with values of the pK_{rh} of 5.5 and 5.3, respectively. The Hot75m1 H77K and Hot75m1 H77D mutants, where the histidine is changed to lysine or aspartic acid, respectively, show no or very little pH dependent change in spectral properties

and appear to be only in the basic form in the pH range tested here. It is likely caused by the pK_{rh} value being shifted so much to be outside the pH range of 4.9-9.1.

Thus, we have shown that mutations in the conserved histidine residue (H75 in Bac31A8 and H77 in Hot75m1) in two different proteorhodopsin variants give the same results. When the histidine is changed to an asparagine, glutamine, glutamic acid or a tryptophan, the value of the pK_{rh} of the spectral shift was significantly lower. A proteorhodopsin with the histidine changed to lysine or aspartic acid shows no significant titration in the pH range tested, and appears to be in the basic form in the pH range of 4.9-9.1.

(b) Arginine Residue

A conserved arginine residue (R94 in Bac31A8 and R96 in Hot75m1) may affect the protonatable group or groups responsible for the pH dependent change in the spectral properties of proteorhodopsin. Therefore, the R96 amino acid residue in Hot75m1 was mutagenized to alanine, glutamic acid or glutamine, the proteorhodopsin variants were expressed, their spectra in intact cells at different pH values were measured and the values of the pK_{rh}'s of the spectral shift were calculated as described previously.

The spectra of Hot75m1 R96A, Hot75m1 R96E and Hot75m1 R96Q at different pH were determined.

Table 7 summarizes the pK_{rh} values and the absorption maxima of the acidic and basic spectra of the different variants of Hot75m1.

TABLE 7. pK_{rh} and Absorption Maxima of Hot75m1 Mutants

Protein	pK_{rh}	Basic Max (nm)	Acidic Max (nm)
Hot75m1	8.2	493	546
Hot75m1 R96A	8.0	492	512
Hot75m1 R96E	7.0	494	509
Hot75m1 R96Q	8.7	493	507

All three R96 mutants showed a significant change in wavelength maximum of the acidic form, with absorption maximum values in the 507-512 nm range compared with an acidic form wavelength maximum of 546 nm observed with the wild-type Hot75m1 protein.

Thus, the Hot75m1 R96 mutants have a smaller difference in absorption maxima of the acidic and basic forms than the wild-type Hot75m1 proteorhodopsin. Such mutants are useful in some technical applications to provide a narrow bandwidth of the absorbance peak at a pH close to the pK_{rh} in a device.

Mutations in the conserved R96 residue of Hot75m1 resulted in the pK_{rh} values that were slightly different from that observed with the wildtype Hot75m1 protein. The Hot75m1 R96A and Hot75m1 R96E mutants have pK_{rh} values of 8.0 and 7.0, respectively, slightly lower than the that of the wild-type Hot75m1 protein. The Hot75m1 R96Q protein where the arginine is changed to glutamine has a value of the pK_{rh} of 8.7, slightly higher than that of the wildtype Hot75m1 protein. Thus, mutations in the R96 residue of Hot75m1 have some effects on the value of the pK_{rh} of the spectral shift.

EXAMPLE 3. Purification of Proteorhodopsins

The results described above were obtained with proteorhodopsins in intact cells, where the protein is present in the cytoplasmic membrane. To characterize the spectral properties of proteorhodopsin variants in the absence of other cell components, we purified wild-type and different mutant Bac31A8 and Hot75m1 proteorhodopsins as described below.

The proteorhodopsin (wild-type or mutant) expression plasmids were transformed into competent cells of the strain BL21-Codonplus-RIL (Stratagene) as described by the manufacturer. The transformed cells were plated on LA + 0.5% glucose + 100 μg/ml carbenicillin + 25 μg/ml chloramphenicol plates and incubated overnight at 37°C. Cells from these plates were grown in 200 ml LB + 0.5% glucose + 100 μg/ml carbenicillin + 10 μM all-trans-retinal medium at 37°C in baffled 500 ml shakeflasks. The cultures were inoculated directly from fresh transformation plates. Proteorhodopsin expression in the different cultures were induced when the OD_{600} of the cultures were approximately 0.8 (after 4 hours growth) by adding 0.5 mM IPTG + 10 μM all-trans-retinal and the cultures were incubated for additional 4 hours at 37°C. The cells were harvested by centrifugation at 3,500xg for 10 minutes and stored at -80°C. Cells were resuspended in lysis buffer containing 40 mM MOPS pH 7.0, 20 mM $MgCl_2$, 0.2 mg/ml lysozyme, 0.2 mg/ml DnaseI, 2% dodecyl-β-D-maltoside, and protease inhibitors (Complete, EDTA-free Protease Inhibitor Cocktail Tablets from Roche Applied Science; and lysed at 0°C by sonication. The lysates were incubated 16 hours with the resin in Talon spin columns (Clontech) that had been equilibrated with 1 ml wash

buffer containing 40 mM MOPS, pH 7.0 and 0.5% dodecyl-β-D-maltoside. The resin was washed three times with 1 ml wash buffer containing 40 mM MOPS, pH 7.0 and 0.5% dodecyl-β-D-maltoside. Proteorhodopsins were eluted from the resin two times with 0.5 ml elution buffer containing 40 mM MOPS pH 7.0, 0.5% dodecyl-β-D-maltoside and 250 mM EDTA. The two elutates were pooled and EDTA was removed by three successive ten-fold concentrations using a Microcon YM-10 centrifugal filter unit (Millipore) and dilutions with a buffer contained 40 mM MOPS, pH 7.0 as described by the manufacturer. The proteorhodopsin samples were then concentrated ten-fold and stored at 4 °C.

EXAMPLE 4. Measurement of Purified Proteorhodopsin Spectra.

5 µl purified proteorhodopsin was diluted in 500 µl of buffer containing 100 mM either citrate (pH 3.54, 3.97, 4.50, or 5.03), acetate (pH 4.88), MES (pH 5.34, 5.80, or 6.35), MOPS (pH 6.75 or 7.32), TAPS (pH 7.91 or 8.43), CHES (pH 9.00 or 9.50), or CAPS (10.12 or 10.64). Wavelength spectra from 250 nm to 650 nm were obtained on a Cary3 spectrophotometer (Varian). The value of the pK_{rh} of the spectral shift was calculated using equation 1 as described previously.

The spectra of detergent-solubilized and purified wild-type Bac31A8, Bac31A8 H75Q, Bac31A8 H75N and Bac31A8 H75K at different pH values were determined. Table 8 summarizes the pK_{rh} value and the maxima of the acidic and basic absorption spectra of wild-type and different mutants.

TABLE 8. pK_{rh} and Absorption Maxima of BAC31A8 Mutants

Protein	pK_{rh}	Basic Max (nm)	Acidic Max (nm)
Bac31A8	7.1	517	543
Bac31A8 H75Q	5.5	517	551
Bac31A8 H75N	5.2	517	534
Bac31A8 H75K	NA	516	-

Note: "NA" means that no significant pH titration is observed, so a pK_{rh} value cannot be determined. "–" means that only a single absorption maximum can be reported because of the lack of titration in the pH range tested.

These results obtained with detergent-solubilized and purified Bac31A8 are similar to the results obtained with proteorhodopsins in intact cells. When the conserved H75 histidine residue of Bac31A8 is changed to asparagine or glutamine, the values of the pK_{rh} of the spectral shift was significantly reduced. Purified Bac31A8 H75Q mutant has a pK_{rh} value of 5.5 and purified Bac31A8 H75N mutant has a pK_{rh} value of 5.2. In comparison, the purified wild-type Bac31A8 protein has a pK_{rh} value of 7.1. The purified Bac31A8 H75K mutant, where the histidine is changed to lysine, shows no or very little pH dependent change in spectral properties, the protein appears to be locked in the basic form in the pH (pH 4.9-9.1) range tested here.

The spectra of purified wild-type Hot75m1, Hot75m1 H77Q, Hot75m1 H77N and Hot75m1 H77K at different pH values were determined.

Table 9 summarizes the pK_{rh} value and the maxima of the acidic and basic absorption spectra of different variants.

TABLE 9. pK_{rh} and Absorption Maxima of Hot75M1 MUTANTS

Protein	pK_{rh}	Basic Max (nm)	Acidic Max (nm)
Hot75M1	8.4	504	538
Hot75M1 H77Q	6.3	508	531
Hot75M1 H77N	6.5	496	523
Hot75M1 H77K	NA	521	-

Note: "NA" means that no significant pH titration is observed, so a pK_{rh} value cannot be determined. "–" means that only a single absorption maximum can be reported because of the lack of titration in the pH range tested.

These results with purified wild-type and mutant Hot75m1 proteorhodopsins show that when the conserved H77 histidine residue of Hot75m1 is changed to asparagine or glutamine, the pK_{rh} value was significantly shifted towards the acidic pH. Purified Hot75m1

H77Q has a value of the pK_{rh} of 6.3 and purified Hot75m1 H77N has a pK_{rh} value of 6.5, compared with 8.4 of the wild-type protein. The purified Hot75m1 H77K mutant where the histidine is changed to lysine shows no or very little pH dependent change in spectral properties in the pH range (pH 4.9-9.1) tested here, and appears to have only one basic form.

In conclusion, mutations in a conserved histidine residue (H75 in Bac31A8 and H77 in Hot75m1) in two different proteorhodopsin variants give similar results. This was observed both when the protein is present in the cytoplasmic membrane in intact cells and with detergent-solubilized and purified protein. Changing the histidine to an asparagine or a glutamine, results in a significant lower pK_{rh} value. Changing the histidine to a lysine results in a basic form only at pH 4.9-9.1, and likely to have a very low pK_{rh} value.

EXAMPLE 5. Determination of Proton Pumping.

The basic form of proteorhodopsin can undergo a photocycle that includes the formation of the excited M-state (or M-like state) and pumping of a proton across the proteorhodopsin-containing membrane from inside the cell to outside the cell. To examine whether the proteorhodopsin variants do undergo a productive photocycle, we have determined the ability of these variants to pump protons. Proton pumping was measured as a change of the extracellular pH with illumination of cells expressing proteorhodopsin, caused by pumping of protons from inside the cells to the extracellular medium. We have also determined the pH dependence of proton pumping for Bac31A8 wild-type and the H75N mutant.

Production and Preparation of Cells. The proteorhodopsin (wild-type or mutant) expression plasmids were transformed into competent cells of the strain BL21-Codonplus-RIL (Stratagene) as described by the manufacturer. The transformed cells were plated on LA + 0.5% glucose + 100 μg/ml carbenicillin + 25 μg/ml chloramphenicol plates and incubated overnight at 37°C. Cells from these plates were grown in 50 ml LB + 0.5% glucose + 100 μg/ml carbenicillin + 10 μM all-trans-retinal medium at 37°C in baffled 250 ml shakeflasks. The cultures were inoculated directly from fresh transformation plates. Proteorhodopsin expression in different cultures was induced when the OD_{600} of the cultures was approximately 0.8 (after 4 hours growth) by adding 0.1 mM IPTG + 10 μM all-trans-retinal,

and the cultures were incubated for additional 4 hours at 37°C. The cells were harvested by centrifugation at 3,500xg for 10 minutes and stored at -80°C.

Cells were thawed and suspended in 150 mM NaCl without buffer (sparged with helium to remove carbon dioxide). Cells were washed an additional three times with 150 mM NaCl without buffer. The cells were suspended to an OD_{600} of 2.0 in 150 mM NaCl. Cell samples were adjusted to between pH 7.5 and 8.5 with 0.5 M HCl or 0.5% NaOH.

Equipment and methods. The values of pH of 2.0 ml cell samples in 10 ml borosilicate glass tubes were measured using a Beckman (Fullerton, CA) Phi40 pH meter with a voltage-out port connected to a National Instruments (Austin, TX) data acquisition board, which averaged the signal at one-second intervals and saved the data on a Pentium based personal computer. Cell samples were magnetically stirred and were illuminated from two sides by split fiber-optic cables attached to a 150-watt illuminator (Fiber-Lite A3200) from Dolan-Jenner Industries (Lawrence, MA) with an infrared filter (03FC569) from Melles Griot (Irvine, CA) and a 450 nm cut-on colored-glass filter (51284) from Oriel Instruments (Stratford, CT). The manual opening and closing of a shutter controlled light. The incident angle of light was held constant by the use of a custom-built cell holder. After the determination of pH for 3 minutes prior to illumination, the pH of the cell sample was determined during the 3 minutes of constant illumination. The pH of the cell sample was then measured for 3 minutes after illumination was stopped. This process was repeated two more times for each cell sample resulting in three repeated measurements of each cell sample.

FIGURE 7-1 shows that the pH of the cells expressing Bac31A8 wild-type, which is an active proton pump, drifted slowly over time without illumination. However, a rapid (within a few seconds) change in pH occurred following illumination. After illumination was ended, the pH value of the cell samples returns to that before illumination. This pH change is a result of cellular mechanisms that control membrane potential.

FIGURE 7-2 shows that the pH of cells expressing a LacZ control protein also drifted slowly over time without illumination. With illumination, the pH shifted slowly (over the entire 3 minutes) and minimally (at most 0.05 pH units), presumably from a temperature change at the pH probe. This temperature effect was reduced dramatically by the addition of the infrared filter, but it could not be completely eliminated.

Data Analysis. To remove the drift of the extracellular pH values with time mathematically, the data not affected by illumination was fitted to a model-independent polynomial approximation, equation 2,

$$pH = A \times t^4 + B \times t^3 + C \times t^2 + D \times t + E \quad (2)$$

where pH is the measured value of pH when illumination no longer affected pH (i.e. last ten seconds of data collected without illumination), A, B, C, D and E are empirically fit parameters and t is the time at which the values of pH were measured. The results of the above equation were subtracted from the entire data set to give the illumination dependent changes in pH. The illumination dependent changes were fitted to equation 3, an exponential decay model assuming that cells dynamically adjust proton flow to maintain membrane pH differential,

$$\Delta pH = \Delta pH_{max} - \Delta pH_{max} \times e^{-kt} \quad (3)$$

where ΔpH is the change in pH due to illumination, ΔpH_{max} is a fit parameter indicating the maximal change in pH, k is a fit parameter indicating the rate at which that maximum is reached and t is the time after illumination is initiated. The value of the product of the two fit parameters, $\Delta pH_{max} \times k$, is the slope of a line tangent at $t = 0$. This value represents the proton-pumping rate in units of ΔpH / min before a cellular response to proton pumping.

Ability to Pump Protons. To determine whether a proteorhodopsin variant is capable of pumping protons, the value of $\Delta pH_{max} \times k$ for cells expressing a proteorhodopsin variant is compared using the student-t test to the value of $\Delta pH_{max} \times k$ determined for cells expressing a LacZ control. The degrees of freedom, n, and the results of the student-t test at 60-99% confidence are given in Table 10. This test indicates whether a particular proteorhodopsin mutant is a proton pump at a given confidence. The negative result of this test does not indicate that a given proteorhodopsin mutant is not a proton pump; it only indicates that the assay is not sensitive enough to detect pumping by that mutant.

TABLE 10. Proton Pumping Results. (Result of t-test at given Confidence Interval)

Proteorhodopsin	n	60%	75%	90%	95%	97.5%	99%
Bac31A8 Wild-type	19	Pump	Pump	Pump	Pump	Pump	Pump
Bac31A8 D97N	19	Not detected	Not detected	Not detected	Not detected	Not detected	Not detected
Bac31A8 E108Q	19	Not detected	Not detected	Not detected	Not detected	Not detected	Not detected
Bac31A8 H75K	28	Pump	Pump	Pump	Not detected	Not detected	Not detected
Bac31A8 H75K [a]	40	Pump	Pump	Pump	Pump	Pump	Not detected
Bac31A8 H75N	19	Pump	Pump	Pump	Pump	Pump	Pump
Bac31A8 H75Q	16	Pump	Pump	Pump	Pump	Pump	Pump
Hot75m1 Wild-type	22	Pump	Pump	Pump	Pump	Pump	Pump
Hot75m1 H77K	19	Pump	Not detected	Not detected	Not detected	Not detected	Not detected
Hot75m1 H77K [a]	31	Pump	Not detected	Not detected	Not detected	Not detected	Not detected
Hot75m1 H77N	16	Pump	Pump	Pump	Pump	Pump	Pump
Hot75m1 H77Q	16	Pump	Pump	Pump	Pump	Pump	Pump
Hot75m1 H77D	22	Not detected	Not detected	Not detected	Not detected	Not detected	Not detected
Hot75m1 H77D [a]	34	Not detected	Not detected	Not detected	Not detected	Not detected	Not detected
Hot75m1 H77E	19	Pump	Pump	Pump	Pump	Pump	Pump
Hot75m1 H77W	16	Pump	Pump	Pump	Pump	Pump	Pump
Hot75m1 R96A	16	Pump	Pump	Pump	Pump	Pump	Pump
Hot75m1 R96E	22	Pump	Pump	Pump	Not detected	Not detected	Not detected
Hot75m1 R96E [a]	34	Pump	Pump	Pump	Pump	Not detected	Not detected
Hot75m1 R96Q	19	Pump	Pump	Pump	Pump	Pump	Pump

Note: "Pump" indicates that proton pumping is detected at the respective confidence level. "Not detected" indicates that no strong evidence of proton pumping exists, but does not indicate that proton pumping is not occurring. [a] The data for the two proteins with mutations at residues central to the proton pumping mechanism, Bac31A8 D97N and Bac31A8 E108Q, were included with the non proton pumping control data to allow for a greater value of n and a greater discrimination of proton pumping capability.

Using this assay, we observe proton pumping by the Bac31A8 and Hot75m1 wild-type proteorhodopsins as expected. No detectable proton pumping was observed with the Bac31A8 D97N mutant proteorhodopsin, which appears to exist in an acidic form only (Dioumaev, *et al.*). Since proteorhodopsin in the acidic form is unable to pump protons (Dioumaev, *et al.*), this confirms the feasibility to use this assay for determining whether proteorhodopsin mutants can undergo a photocycle that includes proton pumping. Proton pumping by the Bac31A8 E108Q, where the photocycle is more than a hundred fold slower than that of the wild-type protein mutant, was not detected. Proton pumping was detected for Bac31A8 wild-type, H75K, H75N, and H75Q; and Hot75m1 wild-type, R96A, H77K, H77N, H77Q, H77E, H77W, R96E and R96Q; which indicates that these proteins, when exposed to light, undergo a productive photocycle that includes proton pumping.

Dependence of Extracellular pH on Proton Pumping. The proteorhodopsins with mutations in the conserved histidine residue are expected to be better proton pumps at low pH than wildtype proteorhodopsin, since the lower pK_{rh} value of the mutants results in a larger fraction of the protein in a productive basic form at low pH. To confirm this, we have compared the rate of proton pumping by Bac31A8 wildtype and the H75N mutant.

Cell samples produced in a manner similar to those used to detect proton-pumping activity were adjusted with 0.5 M HCl or 0.5% NaOH to various pH values and the rate of pumping was determined. The value of extracellular pH was determined from the average pH value over the course of the experiment. FIGURE 7-3 displays the dependence of proton-pumping rate on the extracellular pH for the Bac31A8 wild-type and the H75N mutant. The results show that the Bac31A8 H75N mutant is a better proton pump than the Bac31A8 wild-type at acidic pH. The results confirm that a mutation lowering the pK_{rh} value results in a proteorhodopsin mutant having an extended pH range where a productive photocycle is present.

Although the invention has been described with reference to the presently preferred embodiments, it should be understood that various modifications can be made without departing from the scope of the invention.